70-24,498

VACHTSEVANOS, George John, 1938-
CONDITIONAL OPTIMIZATION OF NONLINEAR
ADAPTIVE CONTROL SYSTEMS.

The City University of New York, Ph.D.,
1970
Engineering, electrical

Conditional Optimization

of

Nonlinear Adaptive Control Systems

by

George J. Vachtsevanos

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1970

This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

April 29 1970
date

Chairman of Examining Committee

24 April 1970
date

Executive Officer

Prof. Vincent Deltoro

Prof. Stanley Katz

Prof. Ralph Mekel

Supervisory Committee

The City University of New York

PLEASE NOTE:

Figures are not original copy.
Light and indistinct print on
several pages. Filmed as
received.

UNIVERSITY MICROFILMS.

To

my wife Athena

and

the memory of my father

ACKNOWLEDGEMENT

The author wishes to express his sincere appreciation to Professor George M. Kranc for his guidance and many contributions to the thesis.

The author is deeply indebted to Professors Ralph Mekel of the Electrical Engineering Department and Stanley Katz of the Chemical Engineering Department for their many suggestions during the completion of the research work. Gratitude is also expressed to Professor Shee-Ming Chen for his valuable assistance in the computer programming of the case studies.

The research was supported in part by NSF Grant No. GP321 and numerical calculations were carried out at the City College Computational Center.

The shades of night were falling fast,
As through an Alpine village passed
A youth, who bore, 'mid snow and ice,
A banner with the strange device,
Excelsior!

In happy homes he saw the light
Of household fires gleam warm and bright;
Above the spectral glaciers shone,
And from his lips escaped a groan,
Excelsior!

"Try not the pass!" the old man said;
"Dark lowers the tempest overhead,
The roaring torrent is deep and wide!"
And loud that clarion voice replied,
Excelsior!

"O stay," the maiden said, "and rest
Thy weary head upon this breast!"
A tear stood in his bright blue eye,
But still he answered with a sigh,
Excelsior!

Beware the pine-tree's withered branch!
Beware the awful avalanche!"
This was the peasant's last Good-night,
A voice replied, far up the height,
Excelsior!

A traveller, by the faithful hound,
Half-buried in the snow was found,
Still grasping in his hand of ice
The banner with the strange device,
Excelsior!

Longfellow

ABSTRACT

The thesis is concerned with a design algotithm for the adaptive control of systems described by nonlinear operator equations.

It is based on Kulikowski's technique as presented for systems described by integral operator equations. The approach is extended to include bounded input functions and the limited range of performance criteria, considered previously, is similarly extended making use of the concepts of nonlinear functional analysis. Systems described by nonlinear differential equations are also considered. In both cases the plant operator is assumed to be only partially known, with no a priori knowledge of either the Volterra Kernels or a vector of slowly varying parameters. The optimization problem is formulated as a conditional-minimization problem and the solution is obtained in the form of an iteration. The identification procedure involves the evaluation of certain essential plant characteristics (measurable differentials); appropriate techniques for the measurement of plant differentials and the estimation of their adjoints are presented. Whenever on-line measurements are not feasible due to input constraint conditions, a plant model is used; procedures for modelling the plant dynamics

are therefore examined. Continuous and discrete, single input-output and multivariable cases are studied. Contraction mapping, Newton's method and Altman's gradient procedure are three iteration techniques used in the overall implementation scheme. Three case studies with their digital computer simulation are also included in the thesis.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The thesis is concerned with the adaptive optimization of systems described by nonlinear integral or differential equations. The plant dynamics are assumed to be only partially known containing a vector of unknown but slowly varying parameters. The space of input functions is assumed to be bounded. The basic optimization alorithm utilized in this thesis was introduced by Kulikowski.

I - 1. The Adaptive Problem.

There is not, as yet, a generally accepted definition of an adaptive control system. Truxal [1] states " An adaptive system is any physical system which has been designed with an adaptive viewpoint." A functional breakdown of an adaptive system proposed by G.R. Cooper, J.E. Gibson et.al. [2] which clearly places the adaptive nature in evidence is the following: The system must a) provide continuous information about the present state of the system or identify the process: b) compare present system performance and make a <u>decision</u> to adapt the system so as to achieve optimum performance, and c) initiate a proper <u>modification</u> so as to drive the control system to the optimum. With this breakdown then in mind an adaptive control system is defined here as a control system which is capable of monitoring its own performance or optimum condition and modifying its behavior in such a manner as to optimize the index of performance or approach the optimum condition. Y.Z. Tsypkin [3] signaling the beginning of a new period in automatic control theory, the period of adaptation, writes: " The possibility of controling objects with incomplete a priori information (or even no a priori information at all) is based on application of adaptation and sel-organization in automatic systems that decrease the initial indeterminacy by using information obtained during the course of the control process."

There is no systematic approach to the problem of adaptive control. Many of the pioneering schemes had been proposed with a particular application in mind. Draper and Li's [4] system identifies the index of performance (which establishes a quality level of system response) and a sinusoidal perturbation input generates changes in the index of performance value containing a component at the perturbing frequency. This component after processing is fed back as a correction signal to change an adaptive parameter, which in turn maintains the value of the IP (Index of Performance) within specified bounds. Tsien's [5] approach and the principle of peak holding is similar in concept to the work of Draper and Li. Here the system continuously hunts back and forth in the near vicinity of the peak IP value. The system of Anderson et.al.[6] varies one independent parameter to hold the IP constant. The identification involves the determination of the coefficients of its describing equation utilizing correlation techniques. Gibson and Meditch [7] investigated the signal synthesis approach to the adaptive control problem. The input signal is synthesized over discrete intervals of time on the basis of the system's past history and its predicted desired performance over the next time interval. Whitaker et.al. introduced the model reference adaptive system and some of the first contributions in the optimal-adaptive area were given by Kalman, Merriam, Braun and Hsieh. An extensive review of adaptive technology up to 1965 is presented in references 8 and 9.

Three categories which appear to encompass a large proportion of adaptive control systems are (1) high-gain schemes, (2) model-referenced schemes and (3) optimum adaptive schemes. The first has found wide applications and is based on keeping the gain in the feedback loop around the changing process as high as possible so that the input-output transfer is kept close to unity. It requires the presence of a model and the knowledge of a large amount of a priori information about the process because of stability problems arising at high gain.

The model reference adaptive scheme employs a reference model which is an analog or digital representation of the desired dynamic response of the system. In one configuration the actual process output is compared to the model output while both are subjected to the same input and an adjusting mechanism changes the adaptive parameter of the controller according to some criterion on the error function. Mathematical programing techniques or Lyapunov's direct method are utilized for the functional description of the adjusting mechanism. A second approach to model reference adaptive control involves the production of an augmented control signal by the adaptive controller which will cause the output of the process to be the same as the output of the reference model. The controller design is based on stability criteria derived from Lyapunov's direct method.

The general philosophy of optimal-adaptive methods is the following: a control signal is synthesized by solving some optimization problem on the assumption that the process parameters and states are known. The overall scheme then involves state estimation and parameter identification procedures so that an expression for the control signal containing only known quantities be made possible. The adaptive algorithm described in the thesis belongs to this general class of optimal-adaptive schemes.

I - 2. Kulikowski's Approach to the Adaptive Problem.

In 1961 Kulikowski [10, 11] suggested an approach to the problem of the adaptive optimal control of non-linear systems whose dynamic characteristics are only partially known and which vary slowly in time. A performance functional of the form

$$P[x] = \int_0^T G(x,y,y^d)dt \qquad (I-1)$$

and an unbounded set of control functions X are assumed to be given. The minimal value achieved by $P[x]$ on the set X connotes optimal system performance. T is chosen to be equal to the plant settling time for reasons that will be explained shortly, and G is assumed to be twice differentiable in each of its arguments x and y, the differentials taken in the Frechet or strong sense. The over-all system configuration suggested by Kulikowski is shown as Figure I-1. The input to the plant is $x(t)$, $y(t)$ is the plant output and $y^d(t)$ is the desired output value taken to be constant. It is further assumed that:

The output-input relation of the plant can be described by a nonlinear integral operator $A[x]$. This operator, for example, can be of the polynomial type with kernels generally unknown to the controller.

$A[x]$ is twice differentiable.

The sets X and Y of input and output functions respectively are subsets of $L^2(0,T)$ (Hilbert) space.

$G(x,y,y^d)$ is chosen such that Equation (I-1) evaluates the output behavior of the plant over one period of steady state operation. The object is to construct a periodic input signal which forces the plant into a steady state operation and minimizes (or maximizes) the performance functional, Equation (I-1).

Kulikowski's approach involves the construction of a sequence of input functions $x_n(t)$ $n=1,2,\ldots$ in the space X by alternating periods of identification with periods of optimization. An explicit solution for the optimal control signal necessitates the identification of the unknown plant characteristics. At this step of the adaptive problem lies the attractive feature of the approach. The entire plant operator need not be identified but only essential plant characteristics (measurable differentials). The optimization problem is formulated and a solution is obtained using classical Calculus of Variations techniques. The step by step optimization and identification procedure is designed to achieve optimum performance in the limit as $n \to \infty$. Figure I-2 shows on a real time basis the intervals under consideration. During the first interval (o,t_1) necessary information regarding the plant structure and the values of fixed coefficients is obtained by experimentation as it is explained in the next section. It is assumed that the plant dynamics remain fixed during the time required for implementation of the iterative optimization procedure, i.e. no changes of the kernels or plant parameters occur in the interval (0,T').

I - 3. The Identification Problem.

Pearson and Sarachik [12, 13] have extended the Kulikowski approach to include a wide class of systems by reformulating the identification problem to account for certain physical considerations associated with the plant dynamics. It was recognized that the plant operator A which maps an input element x into an output element y is in general not unique. During an interval of time (0,T) the plant operator depends not only on the input applied during the same interval but also on past inputs. In the adaptive procedure described above a step of optimization is accompanied by a period of identification and it is essential that identification and optimization be performed during periods when past inputs have a systematic effect on the output. If the plant possesses no pure integrators within its structure it is possible to accomplish this by preceding each period of identification with a long enough interval of time during which no input is applied and the plant transients are allowed to decay to zero. In general, with plants possessing pure integrators, an output steady state is reached by applying a suitable test periodic input. It is assumed that the system is asymptotically stable with no jump resonances and therefore a periodic input will result in a periodic output of the same period. The form and period of the test input depend on the plant type, i.e. whether the plant possesses none, one, two, etc. pure integrators, and also on the nature of the nonlinearity i.e. whether even or odd. Experimentation with suitable test signals during the learning period $(0,t_1)$ will determine the plant type and nature of the nonlinearity together with values of fixed gains and parameters associated with the linear and nonlinear parts of the system. The latter information is utilized in one specific method for the identification of the plant adjoint differential. Details of the technique can be found in the above references.

The identification procedure is therefore strongly dependent on the plant type and the nature of the nonlinearity.

For type zero plants if a waiting period of length T, with T equal to or greater than the plant settling time, preceded the application of any input element all transients due to previous inputs will decay to negligible values. A second method for systematizing the effect of past inputs involves a single repetition of the input element $x(t)$, $t \in (0,T)$, thus forcing the plant into a steady state operation; T must again be greater than or or equal to the plant settling time.

For plants possessing one pure integrator consider the system shown in Figure I-3. N_1 and N_2 are zero memory odd nonlinearities. The odd property of the nonlinear parts of the system insures preservation of the sign of the input signal at the output terminals. Given any input $\bar{x}$, over $0 < t < T$, a periodic input x can be constructed by setting

$$x(t) = (-1)^{i+1}\,\bar{x}(t + iT) \tag{I-2}$$

$$(i = 0, \pm 1, \ldots, -\infty < t < +\infty)$$

and a periodic output y(t) eventually results with a period of length 2T. It was shown by Pearson that the steady state output is uniquely characterized in this case by the periodic input x(t) as defined above and the integrator accumulation due to portion of the input applied just prior to t = 0; the latter is systematized to affect all the succeeding outputs in the same manner.

If the plant has more than one pure integrator in its structure, the periodic input is constructed similarly with the period of the input and output signals increasing as the number of pure integrators present increases.

I - 4. D.C. Level of System Output.

In some cases system performance depends strongly on the D.C. level of the output y(t). The periodic input for a type one system constructed according to the rule of the previous section can not lead to an optimum because although it establishes proper steady state behavior it does not alter the D.C. level of the system output for any choice of x(t). Pearson [14] introduced as part of the optimization a quantity S(t) called the input accumulation and defined by:

$$S(t) \triangleq \int_{-\infty}^{t} x(\tau)d\tau \tag{I-3}$$

The value of S(t) together with x(t) is chosen at each optimization stage to improve system performance. The D.C. level of the system output is altered by adjusting S(t) through the addition of a fixed quantity d to the first segment of length T of the periodic input i.e.

$$x_i(t) = \begin{cases} x(t) + d & \text{for } i = 1 \\ (-1)^{i+1}\, x(t) & \text{for } i > 1 \end{cases} \qquad t \in (0,T) \tag{I-4}$$

I - 5. Scope and Contributions of the Thesis.

It is the purpose of this thesis to demonstrate the theoretical aspects of a design approach to the adaptive optimal control problem and the applicability of the design algorithm to practical engineering situations.

Chapter II introduces the fundamental tools of functional analysis which are found to be necessary for the formulation and solution of the adaptive problem. The definitions and theorems are drawn here from standard textbooks on functional analysis.

Chapter III deals with plants described by integral operator equations. The main contributions here are:

a) The extension of the Kulikowski approach to plants whose input-output time functions are elements of a general $L^p(0,T)$ - space. This makes possible the inclusion of a wider class of performance criteria which describe the quality of system performance.

b) Certain conditions regarding the properties of the plant differential have been relaxed, i.e. it is found that the linearity of the differential with respect to the input variation is guaranteed even when the plant differential exists in its weak sense only.

c) The conditional optimization problem is solved. A bound is placed on the control function and a Lagrange-multiplier method is used to formulate the augmented performance functional. The solution is effected by means of classical calculus of variations techniques.

d) The utilization of the Liusternik and Weierstrass theorems for the solution of the adaptive problem with input constraints, and

e) the overall implementation scheme.

Chapter IV introduces the conditional optimization problem for plants described by nonlinear differential equations. Most practical processes are described by such equations and the formulation and solution of the identification and optimization problems shed additional light on the physical aspects of the adaptive mechanism. They bring out clearly the major assumptions imposed on the mathematical plant structure and their implications with regard to accuracy of the solution. This approach is one of the main contributions of the thesis.

Application of the adaptive algorithm to complex processes necessitates the use of a digital computer in the control loop. Chapter V develops the discrete single input-output and multivariable versions of the adaptive problem.

Finally, in Chapter VI three case studies are presented. They illustrate some basic concepts of all the different approaches discussed in the thesis. The computational difficulties associated with the implementation of the search algorithms limit severely at the present time the range and complexity of the examples.

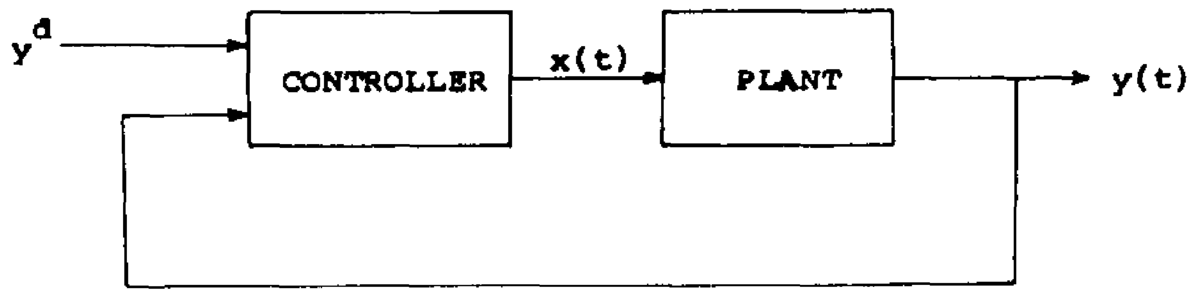


FIGURE I-1

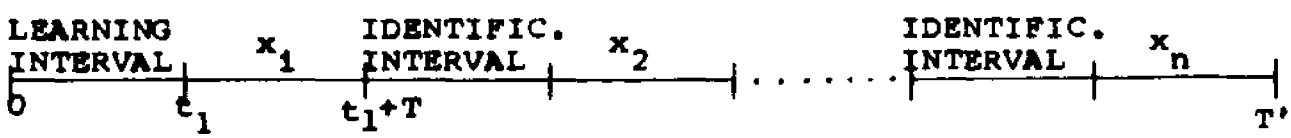


FIGURE I-2

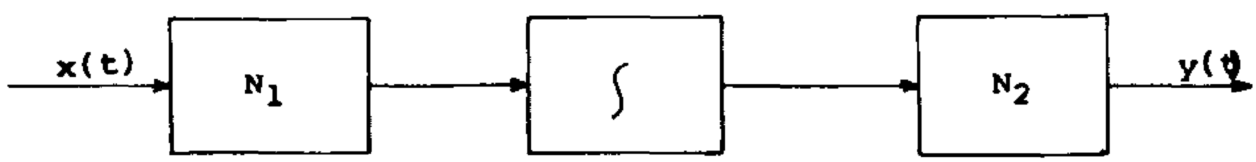


FIGURE I-3

BIBLIOGRAPHY

I-1 J. G. Truxal, "Trends in Adaptive Control Systems," Proceedings of the National Electronics Conference, Vol. 15, 1959.

I-2 G. R. Cooper, J. E. Gibson et. al. "A Survey of the Philosophy and State of the Art of Adaptive Systems" Technical Report No. 1, Contract AF 33(616) - 6890, Purdue University, July 1960.

I-3 Y. Z. Tsypkin, "Adaptation, Training and Self-Organization in Automatic Systems" Automation and Remote Control, Vol. 27, No. 1, Jan. 1966, pp. 16-51.

I-4 L. S. Draper and Y. T. Li, "Principles of Optimalizing Control Systems and an Application to the Internal Combustion Engine," ASME, New York, 1951.

I-5 H. S. Tsien and S. Serdengect, "Analysis of Peak Holding Optimalizing Control," Journal of the Aeronautical Sciences, August 1955.

I-6 G. W. Anderson, Q. N. Buland and G. R. Cooper, "The Aeronutronic Self Optimizing Automatic Control System," Proc. of the Self Adaptive Flight Control

Symposium, WADC Report No. 59-49 WADC Wright - Patterson AFB, Ohio, March 1959.

I-7 J. E. Gibson and J. S. Meditch, "A Class of Predictive Adaptive Controls" Technical Report No. 2, Contract AF 33 (616) - 6890, Purdue University, February 1961.

I-8 Electronics Laboratory, General Electric Company, "A Survey of Adaptive Control Technology," Syracuse, N.Y., 1962.

I-9 P. D. Donalson and F. H. Kishi, "Review of Adaptive Control Systems Theories and Techniques" in Modern Control Systems Theory, book (edited by C. T. Leondes) McGraw-Hill Book Co. 1965.

I-10 R. Kulikowski, "On Optimization of Time Varying Inertial and Nonlinear Control Systems," Bulletin De L'Academie Polonaises Des Sciences, Serie Des Sciences Techniques, Vol. IX, No. 5, 1961, pp. 477-486.

I-11 R. Kulikowski, "On Optimization of Nonlinear Dynamic Processes with Unknown Parameters," Technical Report 71, School of Engineering, Columbia University, 1962.

I-12 A. E. Pearson and P. E. Sarachik, "On the Formulation of Adaptive Optimal Control Problems," Proc. of JACC, pp. 13-22, 1964.

I-13 A. E. Pearson, "Adaptive Optimal Control of Nonlinear Processes," Trans. ASME, Journ. of Basic Engr. Vol. 86, pp. 151-159, 1964.

I-14 A. E. Pearson, "Aspects of Adaptive Optimal Steady State Control," Proc. of JACC, pp. 783-799, 1966.

CHAPTER II

MATHEMATICAL BACKGROUND

In this chapter some basic concepts of functional analysis are presented. The discussion below is intended to introduce only elements of functional analysis which are used in the thesis. The thesis deals mainly with operators on normed linear spaces; such spaces are therefore defined. The adaptive control problem involves sequences of input functions which converge to an optimal solution; the convergence of sequences in normed linear spaces is being discussed. The identification problem utilizes properties of differential operators in the same spaces; certain aspects of the calculus of differential operators are some of the topics of this chapter. Sources from which the definitions and theorems are taken are the books by Kolmogorov and Fomin [1], Liusternik and Sobolev [2], C. T. Leondes [3], Vainberg [4], and Kantorovich and Akilov [5].

II - 1. The Concepts of Set and Space.

The concept of a set is vaguely general. A set may be defined as a collection of its elements. The elements of an abstract set can be of a most general sort, but for most purposes the elements can be thought of as points, numbers, or functions.

The concept of space refers to a set which is taken to be the universe or set of elements used to compose all the sets under consideration. By making various assumptions or definitions various types of spaces are possible.

A <u>metric space</u> consists of a set X and a non-negative real number $\rho(x,y)$ defined for each pair of elements x and y in the set. The number $\rho(x,y)$ is called the distance between the elements x and y, and the following three conditions called the metric axioms, hold:

1. $\rho(x,y) = 0$ if and only if $x = y$ (identity)
2. $\rho(x,y) = \rho(y,x)$ (symmetry)
3. $\rho(x,y) + \rho(y,z) \geq \rho(x,z)$ (triangle inequality)

Examples of metric spaces are:

1. The set D^n of ordered n - tuples of real numbers $x = (x_1, x_2, \ldots x_n)$ with distance function:

$$\rho(x,y) = \left\{ \sum_{k=1}^{n} (y_k - x_k)^2 \right\}^{\frac{1}{2}} \quad \text{(II-1)}$$

called the Euclidean n - space R^n.

2. The set of all continuous functions $x(t)$ on the closed interval $a \leq t \leq b$ with the distance function

$$\rho(x,y) = \left\{ \int_a^b (x - y)^2 dt \right\}^{\frac{1}{2}} \qquad \text{(II-2)}$$

An element x of a metric space M is called the limit of a sequence of elements $x_1, x_2, \ldots x_n \ldots$ of M, if $\rho(x_n, x) \to 0$ as $n \to \infty$.

<u>Theorem</u>.

If a sequence $\{x_n\}$ of points of a metric space X converges to a point $x \in X$, then every subsequence $\{x_{n_k}\}$ of the sequence $\{x_n\}$ converges to the same point. The proof is trivial. For, if $\rho(x_n, x) < \varepsilon$ for $n \geq n_o(\varepsilon)$, then $\rho(x_{n_k}, x) < \varepsilon$ for $n_k \geq n_o(\varepsilon)$. Also a sequence of points of a metric space can converge to at most one limit. If a sequence $\{x_n\}$ of points of X converges to a point $x \in X$, then it is bounded in the sense that for every fixed point of the space, the set of numbers $\rho(x_n, \theta)$ is bounded, i.e. $\rho(x_n, \theta) < K$ for every x_n and a unique constant K.

A <u>linear space</u> is a set R of elements x,y,z,... for which the following operations are defined:

1. Addition. For each pair of elements x,y there is a unique element $z = x + y$ such that:

a) $x + y = y + x$

b) $x + (y + z) = (x + y) + z$

c) there is an element $0 \in R$, $x + 0 = x$ for all $x \in R$.

d) for each $x \in R$, there is an element $-x \in R, x+(-x)=0$.

2. Scalar multiplication. For $x \in R$ there is an element $ax \in R$:

a) $a(bx) = (ab)x$

b) $1\ x = x$

3. Relation between addition and scalar multiplication:

a) $(a+b)x = ax + bx$

b) $a(x+y) = ax + ay$

A <u>normed linear space</u> is a linear space R with a non-negative number $\|x\|$ associated with each $x \in R$. This number is called the norm of x and must have the properties

1. $\|x\| = 0$ if and only if $x = 0$

2. $\|ax\| = |a| \cdot \|x\|$

3. $\|x+y\| \leq \|x\| + \|y\|$

By using $\rho(x,y) = \|x-y\|$ it is seen that a normed linear space is a metric space. Thus the theorems of convergence which apply to any metric space apply in particular to the normed linear space. Convergence in a normed linear space is called convergence in norm, or strong convergence.

A normed linear space X is conveniently defined by specifying a particular norm on that space. The thesis is principally concerned with equivalence classes of functions which comprise $L^p(0,T)$ - spaces and whose norm is defined by

$$\|x\|_p = \left(\int_0^T |x(t)|^p \, dt \right)^{1/p} \qquad p \geq 1 \qquad \text{(II-3)}$$

with $x(t)$, $t \in (0,T)$, belonging to the space $L^p(0,T)$.

The practical significance of this class of spaces lies with the physical interpretation given to the norm as p takes different values.

For $p = 1$ the space $L^1(0,T)$ defines the set of all functions $x(t)$, $t \in (0,T)$ for which the norm

$$\|x\| = \int_0^T |x(t)| \, dt \qquad \text{(II-4)}$$

exists, i.e. $\|x\| < \infty$.

Equation (II-4) defines the set of all functions with bounded "area" specifically the norm of x(t) with $p = 1$ might signify the total output concentration from a chemical reactor process or the amount of fuel feeding through a valve in the time interval (0,T).

The space $L^2(0,T)$ defines the set of all square integrable functions and the corresponding norm is associated with the "energy" or "power" content of x(t).

For $p = \infty$ the space $L^\infty(0,T)$ defines the set of all continuous functions bounded on the interval (0,T).

Liusternik and Sobolev, cf, [2] p. 10, define a metric in $L^{\infty}(0,1)$ by

$$\|x\|_{p=\infty} = \max_{0 \le t \le 1} |x(t)| \qquad \text{(II-5)}$$

If the interval of the independent variable t is (a,b) then it can be transformed into the interval (0,1) by introducing a new independent variable $\tau = \frac{t-a}{b-a}$.

The spaces $L^1(0,T)$, $L^2(0,T)$ and $L^{\infty}(0,T)$ are of particular interest in the conditional optimization problem. The distance functions, as defined above, are used in order to express analytically a varied class of performance indexes and constraint conditions.

Concerning the relationship between different $L^p(0,T)$ spaces corresponding to different values of p, any function $x(t)$, $t \in (0,T)$ which belongs to $L^{p_1}(0,T)$ also belongs to $L^{p_2}(0,T)$ for $p_2 < p_1$, i.e. the space $L^{p_2}(0,T)$ is a subset of the space $L^{p_1}(0,T)$ for $p_2 \le p_1$.

II - 2. Completeness, Convergence and the Principle of Contraction Mapping.

A sequence $\{x_n\}$ of elements of a metric space X is called a Cauchy sequence or fundamental sequence, if for every $\varepsilon > 0$ there is an index $N(\varepsilon)$ such that
$\rho(x_m, x_n) < \varepsilon$ for $m \ge N(\varepsilon)$ and $n \ge N(\varepsilon)$.

<u>Theorem</u>. If a sequence $\{x_n\}$ converges to a limit x, then it is a fundamental sequence.

<u>Proof</u>. If $\{x_n\}$ converges to x, then for given $\varepsilon>0$ it is possible to find a naturel number $N(\varepsilon)$ such that $\rho(x_n,x)<\frac{\varepsilon}{2}$ for all $n\geq N(\varepsilon)$. Then

$$\rho(x_m, x_n) \leq \rho(x_m, x) + \rho(x_n, x) < \varepsilon$$

for arbitrary $m > N(\varepsilon)$ and $n \geq N(\varepsilon)$. The converse is false for an arbitrary metric space since there exist metric spaces in which there are fundamental sequences which do not converge to a limit in the space.

A complete normed linear space is also called a Banach space. $L^p(0,T)$ belongs to the class of Banach spaces. The completeness of $L^p(0,T)$ is shown in reference [2] pp. 19-20.

As an example of the application of the concept of completeness the so-called principle of contraction mappings is considered. Let R be an arbitrary metric space. A mapping B of the space R into itself is said to be a contraction if there exists a number $\gamma<1$ such that

$$\rho(Bx, By) \leq \gamma\rho(x, y) \tag{II-6}$$

for any two points $x,y \in R$; every contraction mapping is continuous. In fact, if $x_n \rightarrow x$, then, by virtue of (II-6) we also have $Bx_n \rightarrow Bx$.

__Theorem.__ Every contraction mapping defined in a complete metric space R has one and only one fixed point, i.e. the equation $Bx = x$ has one and only one solution.

__Proof.__ Let x_o be an arbitrary point in R. Set $x_1 = Bx_o, x_2 = Bx_1 = B^2x_o$, and in general let $x_n = Bx_{n-1} = B^nx_o$. Then

$$\begin{aligned}\rho(x_n, x_m) = \rho(B^nx_o, B^mx_o) &\leq \gamma^n\rho(x_o, x_{m-n}) \\ &\leq \gamma^n\{\rho(x_o, x_1) + \rho(x_1, x_2) + \dots + \rho(x_{m-n-1}, x_{m-n})\} \\ &\leq \gamma^n\rho(x_o, x_1)\{1 + \gamma + \gamma^2 + \dots + \gamma^{m-n-1}\} \\ &\leq \gamma^n\rho(x_o, x_1)\left\{\frac{1}{1-\gamma}\right\}\end{aligned}$$

Since $\gamma < 1$ this quantity is arbitrarily small for sufficiently large n. Therefore the sequence $\{x_n\}$ is fundamental. Since R is complete $\lim_{n\to\infty} x_n$ exists. Setting $x = \lim_{n\to\infty} x_n$ and by virtue of the continuity of this mapping B, $Bx = B \lim_{n\to\infty} x_n = \lim_{n\to\infty} Bx_n = \lim_{n\to\infty} x_{n+1} = x$. Thus, the existence of a fixed point is proved. For its uniqueness consider $Bx = x$, $By = y$ then $\rho(x,y) \leq \gamma\rho(x,y)$, $\gamma < 1$; this implies that $\rho(x,y) = 0$, i.e. $x = y$.

The principle of contraction mappings can be applied to the proof of the existence and uniqueness of solutions obtained by the method of successive approximations. With its help approximating solutions of algebraic, differential, integral and in general nonlinear operator equations can be found.

II - 3. Operators and Functionals on Normed Linear Spaces.

Let X and Y be two normed linear spaces and let the set $\Omega \epsilon X$. If an element $y = A[x] \epsilon Y$ is associated with each $x \epsilon \Omega$, an operator (or mapping) A is said to be defined on Ω, and is said to transform or map Ω into Y. The set Ω is called the domain of definition of the operator A and the space Y its range. An operator f whose range is a set of real numbers is called a functional.

The following definitions distinguish the most important classes of operators:

1. Let X and Y be metric spaces and A an operator from X into Y. The operator A is said to be continuous at the point $x_o \epsilon X$ if $A[x_n] \to A[x_o]$ as $x_n \to x_o$, $(x_n \epsilon X)$.
If X and Y are normed spaces then A is continuous if for every ε there exists a δ such that

$$\begin{aligned} &\|A[x_n] - A[x_o]\| < \varepsilon \\ \text{when} \quad &\|x_n - x_o\| < \delta \end{aligned} \qquad \text{(II-7)}$$

For all x_n, $x_o \epsilon X$. The norms in (II-7) must be interpreted according to whether the domain X is involved or the range Y. $\|A[x]\|$ refers to the norm in the space Y, while $\|x\|$ refers to the norm in the space X. If A is continuous at every point of a set $E \epsilon X$, we simply say that A is continuous on E.

2. If X and Y are normed spaces an operator A from X into Y is said to be homogeneous if

$$A[\lambda x] = \lambda A[x] \qquad (x \in X) \tag{II-8}$$

A is described as additive if

$$A[x_1 + x_2] = A[x_1] + A[x_2] \qquad (x_1, x_2 \in X) \tag{II-9}$$

3. A is called a linear operator if it is additive and continuous on X.

4. An additive operator A mapping a set Ω of a normed space X into a normed space Y is said to be bounded if there exists a constant C such that, for all $x \in \Omega$:

$$\|A[x]\| \leq C \cdot \|x\| \tag{II-10}$$

A necessary and sufficient condition for an additive operator to be linear is that it be bounded. If C_o is a number defined by:

$$C_o = \sup_{\|x\| \leq 1} \|A[x]\| \tag{II-11}$$

then C_o is called the norm of the linear operator A and is written $\|A\|$. On taking $C = C_o = \|A\|$ in (II-10) we get

$$\|A[x]\| \leq \|A\| \cdot \|x\| \tag{II-12}$$

From the above inequality the norm of A is given equivalently by

$$\|A\| = \sup_{\|x\|=1} \|A[x]\| \qquad \text{(II-13)}$$

If the operator A is a linear functional then its norm becomes

$$\|A\| = \sup_{\|x\|=1} |A[x]| \qquad \text{(II-14)}$$

The simple geometrical significance of the number $\|A\|$ is that it is the upper bound of the expansion coefficients of vectors in the transformation produced by the operator A.

If A and B are two linear operators defined on a normed linear space then the following inequality holds:

$$\|A + B\| \leq \|A\| + \|B\| \qquad \text{(II-15)}$$

II - 4. Conjugate Spaces, Inner Products and Adjoint Operators.

The definition of the norm of a linear functional as given above satisfies all of the requirements found in the definition of a normed linear space (cf. II-1, p. 82). Thus, the totality of all linear functionals on a normed space X itself represents a normed linear space; It is called the conjugate space of X and is denoted by $\overline{X}$.

The general form of a linear functional f(x) on $L^p(0,T)$ space is shown to be (cf. II-2, pp. 102-105):

$$f(x) = \int_0^T x(t)y(t)dt \tag{II-16}$$

where $x(t) \in L^p(0,T)$ and y(t) belongs to the conjugate space of $L^p(0,T)$ denoted by $L^q(0,T)$; p and q are related by

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{II-17}$$

The norm of functional (II-16) as determined by definition (II-14) in conjunction with Holder's inequality is given by

$$F = \left[\int_0^T |f(t)|^q \, dt \right]^{\frac{1}{q}} \tag{II-18}$$

The expression given by the right-hand side of equation (II-16) is a bilinear functional of two variables x and y, i.e. it is linear in x and also in y. Such a bilinear expression is called an inner product between the two elements.

The inner product between two elements $x \in X$, $y \in \bar{X}$ is the bilinear functional associated with X and its conjugate space $\bar{X}$, it is denoted by (x,y) and possesses the following properties:

$$
\begin{aligned}
&1. \quad (x,y) = (y,x) \\
&2. \quad (x_1+x_2,y) = (x_1,y) + (x_2,y) \text{ for all } x_1,x_2 \in X,\ y \in \overline{X} \\
&3. \quad (x\ ,y_1+y_2) = (x,y_1) + (x,y_2) \text{ for all } x\ \ X,\ y_1,y_2\ \ \overline{X} \\
&4. \quad (ax,y) = (x,ay) = a(x,y) \text{ for } a \text{ any real number}
\end{aligned}
\tag{II-19}
$$

The inner product on the conjugate spaces $L^p(0,T)$ and $L^q(0,T)$ where p and q are related by (II-17) is

$$(x,y) = \int_0^T x(t)y(t)dt \tag{II-20}$$

with $x(t) \in L^p(0,T)$ and $y(t) \in L^q(0,T)$ or vice versa.

If p = 2 then q = 2 and $\overline{L}^2(0,T) = L^2(0,T)$ and the spaces are said to be self-conjugate. $L^2(0,T)$ is a real Hilbert space for which the norm is derived from the inner product according to

$$\|x\| = \sqrt{(x,x)} \tag{II-21}$$

If X is a Hilbert space then $x \in X$, and $y \in \overline{X}$ are called <u>orthogonal</u> if

$$(x,y) = 0 \tag{II-22}$$

Generalizing the concept of orthogonality to include elements of normed linear spaces the following definition is considered: The sequences $\{x_n\}$, $x_n \in X$, and

$\{y_n\}$, $y_n \in \overline{X}$ are called biorthogonal if for all n and m

$$(y_n, x_m) = \delta_{nm} \tag{II-23}$$

where δ_{nm} is the Kronecker Symbol ($\delta_{nm} = 1$ for $n = m$ and $\delta_{nm} = 0$ for $n \neq m$). If X is self-conjugate, for instance a Hilbert space, then both sequences belong to X. If $\{x_n\}$ and $\{y_n\}$ coincide, the biorthogonality becomes ordinary orthogonality.

Adjoint Operators. Let the linear operator A map X into Y, where X and Y are linear and normed. Furthermore, let g(y) be a linear functional defined on Y, i.e. $g(y) \in \overline{Y}$, then g(y) is defined for $y = A[x]$ where x is an arbitrary element in X.

Since

$$g(y) = g(A[x]) = f(x)$$

f(x) represents a certain functional defined on X, i.e. $f(x) \in \overline{X}$. It is obvious that f(x) is linear. Hence there corresponds to every functional $g \in \overline{Y}$ a functional $f \in \overline{X}$. The collection of all correspondences thus obtained forms a certain operator defined on $\overline{Y}$ and with range belonging to $\overline{X}$. This operator is designated A^* and is called the adjoint operator of A.

If we use the notation (f,x) for the functional f(x), we obtain (g,A[x]) = (f,x) or

$$(g, A[x]) = (A^*[g], x) \tag{II-24}$$

This relation can be taken as the definition of the adjoint operator.

Some basic properties of adjoint operators are listed below:

1. The adjoint operator of the sum of two linear operators is equal to the sum of the adjoint operators

$$(A + B)^* = A^* + B^* \tag{II-25}$$

2. The adjoint operator of the operator λA, where λ is a scalar multiplier, is equal to the adjoint operator of A multiplied by λ :

$$(\lambda A)^* = \lambda A^* \tag{II-26}$$

3. The adjoint of the identity operator on X is the identity operator on $\overline{X}$:

$$I^* = I \tag{II-27}$$

Theorem. The operator A^*, the adjoint of the linear operator A, which maps the Banach space X into the Banach space Y is also linear and $\|A^*\| = \|A\|$. (For the proof see reference (II-1), p. 103).

As an example consider in $L^2(0,1)$ the Fredholm operator

$$y = Kx(t) = \int_0^1 K(t,s)x(s)ds \tag{II-28}$$

with continuous kernel K(s,t). An arbitrary linear functional f(y) on $L^2(0,1)$ has the form:

$$f(y) = \int_0^1 f(s)y(s)ds \quad f \in L^2(0,1) \tag{II-29}$$

Therefore

$$f(Kx) = \int_0^1 f(s) \left\{ \int_0^1 K(s,t)x(t)dt \right\} ds \tag{II-30}$$

An interchange of the order of integration followed by an interchange of the variables s and t results in

$$f(Kx) = \int_0^1 x(s) \left\{ \int_0^1 K(t,s)f(t)dt \right\} ds = \int_0^1 x(s)g(s)ds \tag{II-31}$$

with

$$g(s) = \int_0^1 K(t,s)f(t)dt \quad (g = K^*f) \tag{II-32}$$

In this case adjoining corresponds to a transposition of arguments in the kernel function.

II - 5. Differentials of Operators.

If, for some $x \in X$ and every $h \in X$,

$$\lim \frac{A[x+\gamma h]-A[x]}{\gamma} = dA_w[x,h] \quad \text{(II-33)}$$

exists, then the operator $dA_w[x,h]$ is called the Gateaux or weak differential of the operator A at the point x in the direction h. The limit is understood here as

$$\lim_{\gamma \to 0} \left\| \frac{1}{\gamma} \left\{ A[x+\gamma h]-A[x] \right\} -dA_w[x,h] \right\| = 0 \quad \text{(II-34)}$$

It follows from the definition that for every real a

$$dA_w[x,ah] = a dA_w[x,h] \quad \text{(II-35)}$$

i.e. $dA_w[x,h]$ is a homogeneous operator in h. However, $dA_w[x,h]$ is not always a linear operator in h. The following theorem gives both necessary and sufficient conditions for the linearity of $dA_w[x,h]$ with respect to h. The proof can be found in reference (II-4) pp. 39-40.

__Theorem.__ In order that the Gateaux differential $dA_w[x_o,h]$ of the operator A be a linear operator in h, it is necessary and sufficient that A satisfy the following conditions:

1. A satisfies a weak Lipschitz condition at the point x_o. We say that the operator A satisfies at the point x_o, a weak Lipschitz condition, if to every unit vector h there corresponds a $\delta(h) > 0$ such that if $|\gamma| < \delta(h)$, then

$$\|A[x_o+\gamma h]-A[x_o]\| \leq C \cdot \|\gamma h\|$$

Where $C > 0$ does not depend upon the vector h.

2. $A[x_o+h_1+h_2] - A[x_o+h_1] - A[x_o+h_2] + A[x_o] = 0$

We define below the strong or Frechet differential of $A[x]$: Let $A[x]$ be an operator defined on some normed linear space, and let

$$\Delta A[h] = A[x+h] - A[x] \tag{II-36}$$

be its increment, corresponding to the increment $h = h(t)$ of the independent variable $x = x(t)$. If x is fixed, $\Delta A[h]$ is in general a nonlinear operator on h. Suppose that

$$\Delta A[h] = dA[h] + \varepsilon \|h\| \tag{II-37}$$

where $dA[h]$ is a linear operator and $\varepsilon \to 0$ as $\|h\| \to 0$, then the operator $A[x]$ is said to be differentiable, and the principal linear part of the increment $\Delta A[h]$, i.e. the linear operator $dA[h]$ which differs from $\Delta A[h]$ by an infinitesimal of order higher than one relative to $\|h\|$ is called the strong or Frechet differential of $A[x]$.

Strictly speaking the increment and the variation of $A[x]$ are operators of two arguments x and h and to emphasize this fact we might write

$$\Delta A[x;h] = dA[x,h] + \varepsilon \cdot \|h\| \tag{II-38}$$

<u>Theorem</u>. If the strong differential exists, then the weak differential exists and $dA[x,h] = dA_w[x,h]$. The converse is not always true; for example the operator A defined by

$$A[x] = \int_0^t |x(\tau)| \, d\tau \tag{II-39}$$

where $x(t) \in L^1(0,T)$, possesses the weak differential

$$dA_w[x,h] = \int_0^t h(\tau) \, \operatorname{sgn}\{x(\tau)\} \, d\tau \tag{II-40}$$

where the sgn(z) is defined by

$$\mathrm{sgn}(z) = \begin{cases} 1 & z > 0 \\ -1 & z < 0 \end{cases} \qquad \text{(II-41)}$$

However the Frechet differential operator $dA[x,h]$ does not exist for the operator (II-39).

The majority of the operators considered in the thesis will be assumed to be twice differentiable in the Frechet sense. However, the analysis will not exclude operators possessing weak differentials only, as long as they satisfy the linearity condition with respect to h. A unifoem notation will be used throughout the thesis and the differential of the operator $A[x]$ will be designated by $dA[x,h]$ whether it be in the Frechet or Gateaux sense.

Suppose that the operator A, from the Banach space X into the Banach space Y, has at the point x the differential $dA[x,h]$. This linear operator, which is an element of the Banach space consisting of every linear operator from X into Y is called the derivative of the operator A at the point x and is denoted by $A'[x]$. Cosequently

$$dA[x,h] = A'[x]h \qquad \text{(II-42)}$$

The concept of the derivative operator is the same in both the Frechet and Gateaux sense.

Higher order differentials of an operator A can be obtained by repeated application of equation (II-33) or from the equivalent expression

$$dA_w[x,h] = \left.\frac{\partial A[x+\gamma h]}{\partial \gamma}\right|_{\gamma=0} \qquad \text{(II-43)}$$

then, the n^{th} order differential of the operator A, when it exists, can be obtained from

$$d^nA[x,h_1,h_2,\ldots h_n] = \left.\frac{\partial^n A[x+\gamma_1 h_1 + \cdots \gamma_n h_n]}{\partial \gamma_1 \partial \gamma_2 \cdots \partial \gamma_n}\right|_{\gamma_1 = \cdots \gamma_n = 0} \qquad \text{(II-44)}$$

with $x, h_1 \ldots h_n \in X$.

The n^{th} order Frechet differential is linear with respect to each of its n variations $h_1, h_2 \ldots h_n$.

II - 6. Extrema.

Let f be a real functional defined on the Banach space X. A point $x_o \in X$ is called an extreme point of f if there is a neighborhood $\delta(x_o)$ on which one of the following inequalities holds:

$$1. \quad f(x) \leq f(x_o) \qquad 2. \quad f(x) \geq f(x_o) \qquad x \in \delta(x_o) \qquad \text{(II-45)}$$

The gradient of a functional f is defined in connection with df(x,h) by its relation to the arbitrary variation h(t) in the inner product form

$$df(x,h) = \int_0^T h(t) \cdot (\text{gradf})dt \qquad \text{(II-46)}$$

Further, a point x_o is called a critical point of the functional f if $\text{gradf}(x_o) = \Theta$, the zero element in $\bar{X}$.

<u>Theorem</u>. Let the functional f be defined on a region w of a Banach space X, and let x_o be an interior point of w, at which f has a linear Gateaux differential. Then, in order that the point x_o be extreme, it is necessary that it be critical, i.e. that $\text{gradf}(x_o) = \Theta$.

<u>Proof</u>. Let h be an arbitrary fixed vector in X. Then $f(x_o + \gamma h)$ is a real function defined in some neighborhood of the point $\gamma = 0$ and having a derivative at $\gamma = 0$. Hence, if x_o is an extreme point of f, then $\gamma = 0$ is an extreme point of the function $f(x_o + \gamma h)$ and therefore

$$\frac{d}{d\gamma} f(x_o + \gamma h)\Big|_{\gamma=0} = 0$$

or $df(x_o,h) = 0$ or finally $(\text{gradf}(x_o), h) = 0$. Since h is an arbitrary vector in X, $\text{gradf}(x_o) = \Theta$, which proves the theorem.

A sufficient condition for the existence of extreme points of certain functionals involves properties of the second differential, for example the following assertion is true:

<u>Theorem</u>. Suppose that the functional f, defined on a Banach space X, has first and second linear Gateaux differentials, and the latter satisfies

$$d^2f(x,h,h) \geq \|h\| \cdot \gamma(\|h\|)$$

where $\gamma(t)$ is a non-negative continuous function defined for $t \geq 0$, such that $\lim_{t \to \infty} \gamma(t) = +\infty$.

Then there exists a point at which f has a <u>relative minimum</u>. For the proof see reference (II-4) p. 80.

Let X be a real Banach space, i.e. a complete normed linear space whose elements can be multiplied by real numbers.

A sequence of elements $\{x_n\}$ converges weakly to x_o if the

$$\lim_{n \to \infty} f(x_n) = f(x_o)$$

holds for every linear bounded functional f(x) defined on X. Strong convergence in the norm sense has already been defined in section 1 of this chapter.

A set w is called weakly closed if it contains all of its weak limit points; i.e. if the sequence $\{x_i\}$ belonging to w, converges weakly to an element x_o, then $x_o \in w$.

A set w is called weakly compact in X, if for every sequence $\{x_n\} \subset w$, there is a subsequence $\{x_{n_k}\}$ which converges weakly to some element $x_o \in X$.

A functional f is weakly continuous at the point x_o if, for every sequence $\{x_n\}$ which converges weakly to x_o, the sequence $\{f(x_n)\}$ converges weakly to $f(x_o)$.

A functional f is called lower (upper) semi-continuous if $x_n \to x$ implies

$$\underline{\lim_{n}}\, f(x_n) \geq f(x) \qquad \left[\overline{\lim_{n}}\, f(x_n) \leq f(x)\right]$$

Let $f_1(x)$ a functional to be minimized under the condition that $f_2(x)$, another functional be smaller or equal to a constant R; x_o is defined to be an ordinary point of $f_2(x) = R$ if

$$\|\mathrm{grad} f_2(x_o)\| > 0 \qquad \text{(II-47)}$$

Also, x_o an element of $f_2(x) = R$, is called a conditionally critical point of the functional $f_1(x)$ with respect to $f_2(x) = R$ if

$$\text{grad} f_1(x_o) = \lambda \, \text{grad} f_2(x_o) \tag{II-48}$$

where λ is a constant multplier.

Liusternik's Theorem. In order that an ordinary point of $f_2(x) = R$ be a conditionally extreme point of the functional $f_1(x)$, it is necessary that this point be a conditionally critical point of $f_1(x)$ with respect to $f_2(x)=R$; i.e. Equation (II-48) is satisfied.+

The most interesting case, from a practical point of view, is to find the extremum of a functional not on the surface $f_2(x) = R$, but in a closed region w, $f_2(x) \leq R$. The following theorem due to Weierstrass gives a sufficient condition for the existence of extreme points with respect to w', the bounary of w.

Weierstrass Theorem. Let $f_1(x)$ be a real functional, defined on a weakly closed and weakly compact set w in a

+ An elementary proof of the theorem for two particular cases can be found on page 97 of the reference (II-4)

Banach space, and having a gradient at every point of w. Then

1. If $f_1(x)$ is weakly continuous in w and the condition

$$\|\operatorname{grad} f_1(x)\| > 0 \tag{II-49}$$

holds at every interior point of w, then $f_1(x)$ assumes its extremum on the boundary w' of w.

2. If $f_1(x)$ is weakly continuous in w and Equation (II-49) holds at all but one interior point of w, then $f_1(x)$ assumes at least one of its extreme points on the boundary w'.

3. If $f_1(x)$ is weakly lower (upper) semi-continuous in w, and Equation (II-49) holds at every interior point of w, then $f_1(x)$ assumes its infimum (supremum) on w'.

BIBLIOGRAPHY

II-1. Kolmogorov, A.N. and Fomin, S.V., " Elements of the Theory of Functions and Functional Analysis, Vol. 1, Metric and Normed Spaces," (book) Graylock Press, Richester, N.Y. 1957.

II-2. Liusternik, L.A. and Sobolev, V.J., " Elements of Functional Analysis," (book) Frederick Ungar Publishing Company, N.Y., 1961.

II-3. Leondes, C.T., " Modern Control Systems Theory," (book) University of California Engineering and Sciences Extension Series, Mc Graw-Hill Book Co., N.Y., 1965.

II-4. Vainberg, M.M., " Variational Methods for the Study of Nonlinear Operators," (book) Holden Day Series in Mathematical Physics, San Francisco, 1964.

II-5. Kantorovich, L.V., and Akilov, G.P., " Functional Analysis in Normed Spaces," (book) The Mac Millan Co., N.Y., 1964.

CHAPTER III

THE CONDITIONAL OPTIMIZATION PROBLEM

In this chapter the formulation and solution of the conditional optimization problem are presented. The discussion is presently restricted to plants described by nonlinear integral operators; in the next chapter we are concerned with plants described by nonlinear differential equations. The adaptive nature of the problem arises from the assumption that the plant dynamics are not completely known. The method of solution to be followed is applicable only when the plant dynamics vary slowly in time. (These general remarks will be amplified and take concrete form shortly in the course of the presentation.) Thus the technique can be applied to complex chemical or manufacturing processes with long enough intervals of operation during which the process dynamics, though partially unknown, remain unchanged and necessary computational aspects of the optimization problem are carried out. The overall adaptive control scheme is designed for on-line operation with a provision for plant modeling and off-line control whenever certain restrictive conditions dictate for this kind of operation as it will be explained in this chapter. Finally, in order to simplify the exposition of basic concepts a one-input, one-output plant is considered first with the discussion of the multivariable case deferred to a later section.

III - 1. Problem Formulation.

It is assumed + that the plant to be controlled is described by the following operator relationship

$$y(t) = A[x(t)] \tag{III-1}$$

where x(t) is the input variable and belongs to a given set X of control signals, y(t) is the output variable and A is generally a nonlinear integral operator. The plant configuration considered is shown in Figure (III-1). The operator A maps elements x from $X \subset L^p(0,T)$-space into an image output space $Y \subset L^\gamma(0,T)$. For example, if the process is structured as shown in Figure (III-2), with N_1 and N_2 zero-memory nonlinear devices and L the linear part, its input-output relation is

$$y(t) = N_2 \left\{ \int_{-\infty}^{t} k(t,\tau) N_1[x(\tau)] d\tau \right\} \tag{III-2}$$

+ A minimal set of basic assumptions is presented in this section; as specific identification and computational techniques are discussed additional assumptions are found to be necessary.

Here the impulsive response of the linear part of the process is considered to be unknown. One of the two methods of identification discussed later in this chapter requires knowledge of the form of the nonlinearities N_1 and N_2, i.e. saturation, cubic, arctan etc., while necessary information concerning fixed parameters associated with N_1 and N_2 are obtained during the initial learning interval; the second identification method is more general in scope and does not require explicit knowledge of the nonlinearities.

More generally, the integral operator formulation can be of the polynomial type:

$$A[x] = A_0(t) + \sum_{i=1}^{\infty} \int_0^t \cdots \int_0^t k_i(t, \tau_1, \ldots \tau_i) x(\tau_1) \ldots x(\tau_i) d\tau_1 \cdots d\tau_i \quad \text{(III-3)}$$

The function $A_0(t)$ represents the initial conditions of the plant and the kernels $k_i(t, \tau_1, \ldots \tau_i)$ are not specified completely.

A quantitative measure of the quality of overall system performance is given in terms of the performance functional $J[x]$. In generall, $J[x]$ is considered to be of the integral form:

$$J[x] = \int_0^T G(x, y, y^d)\, dt \quad \text{(III-4)}$$

where y^d is the desired output state.

The objective is to determine a control signal x which belongs to X and minimizes (or maximizes) $J[x]$, while

$$\|x\|_p \leq R \qquad \text{(III-5)}$$

i.e. the L^p norm of the control signal is constrained to be less or equal to a positive constant R.

The following assumptions are held to be true:

1. $A[x]$ is twice differentiable. The first differential of $A[x]$ will be denoted by $dA[x,h]$, where $h \in X$ is an arbitrary variation of the input variable x. If the integrand of the operator equation defining $A[x]$ is continuous and $A[x]$ is bounded, the differential $dA[x,h]$ is linear with respect to the variation h. A large class of physical systems exhibit such behavior and in the thesis the above linearity condition will be assumed as being satisfied.

2. The process is asymptotically stable. The scheme can be extended to classes of processes which are not asymptotically stable when this condition arises due to the presence of pure integrators in the plant structure. Here knowledge of the number of pure integrators in the linear part of the plant is required. The identification and optimization procedures are

modified according to the work presented by Pearson and Sarachik as described in the first chapter of this thesis. As a result of the stability considerations a periodic input results in a periodic output. The nonlinear devices exhibit no jump discontinuities thus excluding the presence of subharmonic oscillations and assuring that the period of the output is the same as that of the input.

3. Zero-memory nonlinear devices.

4. $G(x, y, y^d)$ is a given positive twice differentiable function in x and y and y^d is the desired output.

5. It is assumed that the dynamic system is optimizable with respect to the given quality functional $J[x]$ defined on the given control signal set X in the sense that $J[x]$ assumes at least one extreme value for $x \in X$. As a counter-example consider the performance criterion:

$$J[x] = \left[\frac{1}{T}\int_0^T |y^d - y(t)|^2 dt \right]^{\frac{1}{2}}$$

with $y^d(t)$ a unit step function, and the plant given by

$$y(t) = \int_0^t x(\tau)d\tau$$

If $X \subset L^2(0,T)$ it is easily shown that the minimum value of $J[x]$ is achieved when $x(t) = \delta(t)$. But $x(t) = \delta(t)$ is an infinite power control signal, i.e. $x(t)$ does not belong to X, therefore the plant is not optimizable. A general enough criterion of optimizability was given by Kulikowski [III-1].

III - 2. General Solution.

The unconstrained problem is considered first and the functional $J[x]$, Equation (III-4), is taken as a measure of system performance.

The necessary and sufficient conditions that $J[x]$ attain a relative minimum or maximum value with respect to the control signal x are that the first differential $dJ[x,h]$ vanish and the second differential $d^2J[x,h,h]$ be strictly positive for a relative minimum or strictly negative for a relative maximum value of the performance functional (see Chapter II, Section 6).

The first differential of $J[x]$ is given by:

$$dJ[x,h] = \int_0^T \left\{ \frac{\partial G}{\partial x} h(t) + \frac{\partial G}{\partial y} dA[x,h] \right\} dt \qquad \text{(III-6)}$$

where $h(t) \in L^p(0,T)$, $dA[x,h] \in L^r(0,T)$

and $\frac{\partial G}{\partial x} \epsilon L^q(0,T)$, $\frac{\partial G}{\partial y} \epsilon L^s(0,T)$

with $\frac{1}{p} + \frac{1}{q} = 1$, $\frac{1}{r} + \frac{1}{s} = 1$

The second differential of the performance functional can be written as

$$d^2J[x,h,h] = \int_0^T \left\{ \frac{\partial^2 G}{\partial x^2} h^2 + 2 \frac{\partial^2 G}{\partial x \partial y} h \cdot A[x,h] + \right.$$

$$\left. + \frac{\partial^2 G}{\partial y^2}(dA[x,h])^2 + \frac{\partial G}{\partial y} d^2A[x,h,h] \right\} dt \tag{III-7}$$

The sufficiency condition, Equation (III-7), for plants with partially unknown dynamic characteristics, is difficult to prove and/or implement and will be considered here as being satisfied; moreover, the problems are physically motivated and, in general, it should be intuitively clear whether the performance criterion is minimized or maximized.

Making use of the definition of the gradient of a functional (section II-6, Equation II-46) Equation (III-6) can be reformulated in the inner product form:

$$dJ[x,h] = \int_0^T h(t) \text{grad} J[x] \, dt \tag{III-8}$$

Since h(t) is an arbitrary variation with respect to x in the interval (0,T) the necessary condition for the existence of an extremum of J[x] can be restated as follows:

$$\mathrm{grad}J[x] = 0 \quad \text{for } t \in (0,T) \tag{III-9}$$

In order to express $dJ[x,h]$ in the form of Equation (III-8) the adjoint differential operator $dA^*[x,g]$ is defined by the inner product relationship:

$$\int_0^T g[x(t)]\, dA[x,h]\, dt = \int_0^T h(t)\, dA^*[x,g]\, dt \tag{III-10}$$

with $g \in \bar{Y} \subset L^s(0,T)$, where $\bar{Y}$ is the space conjugate to Y, $h \in X \subset L^p(0,T)$, and $dA^*[x,g] \in L^q(0,T)$, i.e. dA^* maps elements of L^s (the conjugate of the image space Y) into the space L^q (the conjugate to the domain of the operator A).

Letting $g \triangleq \frac{\partial G}{\partial y}$ and following appropriate substitution in terms of the adjoint differential, Equation (III-6) is written as

$$dJ[x,h] = \int_0^T h(t)\left\{ \frac{\partial G}{\partial x} + dA^*[x,g]\right\} dt \tag{III-11}$$

The gradient of $J[x]$ is therefore recognized to be

$$\mathrm{grad}J[x] = \frac{\partial G}{\partial x} + dA^*[x,g] \tag{III-12}$$

Thus the problem involves the solution of a generally nonlinear operator equation ($\mathrm{grad}J[x] = 0$) in terms of the control signal x(t). For that purpose the unknown adjoint differential $dA^*[x,g]$ is related to measurable plant output differentials; two such identification procedures are presented below. A number of well-known iteration schemes for the solution of the operator equation are also discussed in subsequent sections.

Let us consider next the constraint problem: We seek a minimum (or maximum) of $J[x]$ while an appropriate norm of the control signal is constrained to be less or equal to a positive constant R, i.e.

$$\|x\|_p \leq R \qquad \text{(III-13)}$$

Two distinct possibilities present themselves with respect to the solution of the constraint problem: Either the norm of the control signal is less than R, and therefore the solution is identical to the one considered already for the unconstraint problem, or the norm of x(t) is greater than the value of the constraint. Consider the second case to hold true here. In most control problems whenever the control signal exceeds the constraint boundary the best possible input is found to lie on the boundary of the constraint. By recalling the Weierstrass Theorem (Chapter II, Section 6) operation on the boundary of the constraint is for all practical purposes assured whenever the norm of the gradient of the performance functional does not become identically equal to zero throughout the optimization interval (0,T). We assume operation on the boundary if $\|x\|_p > R$ especially since the problems in this thesis are physically motivated. The inequality constraint becomes an equality, i.e.

$$\|x\|_p = R \tag{III-14}$$

Now, a Lagrangian formulation is used in writing the augmented performance criterion $P[x]$ as

$$P[x] = J[x] + \lambda \|x\|_p \tag{III-15}$$

with the Lagrange-multiplier λ to be evaluated from the constraint condition (III-14).

The necessary condition for the existence of an extremum of $P[x]$, i.e. the vanishing of its first differential, is expressed as

$$dP[x,h] = dJ[x,h] + \lambda \left(\int_0^T |x|^p dt \right)^{\frac{1}{p}-1} \int_0^T |x|^{p-1} \operatorname{sgn} x \cdot h \, dt = 0 \tag{III-16}$$

or equivalently:

$$\operatorname{grad} J[x] + \lambda \cdot \frac{|x|^{p-1} \operatorname{sgn} x}{\|x\|_q} = 0 \tag{III-17}$$

with $\operatorname{grad} J[x]$ given by Equation (III-12).

Thus, utilizing some iterative procedure in conjunction with Equation (III-17), a sequence of input functions can be constructed with each member of the sequence computed from the previous one according to the relationship:

$$x_{k+1} = S(x_k,\ dA^{*}[x_k],\ \lambda) \qquad \text{(III-18)}$$

together with the constraint condition for the determination of the Lagrange-multiplier λ. The limiting element of the sequence constructed from Equation (III-18) results in the optimal input $x^{o}(t)$, $t \in (0,T)$.

The Weierstrass condition can be easily tested in a digital mechanization of a specific process as it will be shown in Chapter VI where applications and their digital implementation are taken up.

In order to illustrate the general solution in some detail for physically meaningful situations let us consider a regulator problem with the performance criterion $J[x]$ of the form:

$$J[x] = \left\| y^{d} - A[x] \right\|_{p_1} \qquad \text{(III-19)}$$

where $y^{d} - A[x]$ is the difference between the actual and desired output values. When $p_1 = 1$ a measure of the integral of the absolute value of the error function is being minimized. With $p_1 = 2$ the criterion is the root mean squared error and for $p_1 = \infty$ the maximum amplitude of the difference between the actual and desired output values is the subject of minimization.

Another form of the performance criterion useful in practical applications for extremum control problems is the following:

$$J[x] = \|A[x]\|_{p_1} \tag{III-20}$$

Here for $p_1 = 1$ the "area" or total output product is maximized; with $p_1 = 2$ we seek an extremum of the output "power" or "energy". In general, forms similar to Equations (III-19) and (III-20) are extremely helpful in translating a great variety of qualitative physical criteria into functional relationships. This clearly demonstrates the usefulness of the extension of the analysis into L^p-spaces.

For purposes of an explicit analytical solution it will be assumed that the performance functional is given in terms of Equation (III-19) with the control signal constrained by

$$\|x\|_{p_2} \leq R \tag{III-21}$$

where p_1, $p_2 \geq 1$ and R is a positive constant. The constraint inequality (III-21) takes on different physical interpretations as p_2 takes on different values in a similar fashion as that described above for $J[x]$.

Following the standard procedure the augmented performance functional $P[x]$ is formed according to the relation:

$$P[x] = \|y^d - A[x]\|_{p_1} + \lambda \|x\|_{p_2} \tag{III-22}$$

where λ is the Lagrange-multiplier and will subsequently be determined from the input constraint condition. $P[x]$ may be written alternatively as:

$$P[x] = \left(\int_0^T |y^d - A[x]|^{p_1} dt \right)^{\frac{1}{p_1}} + \lambda \left(\int_0^T |x|^{p_2} dt \right)^{\frac{1}{p_2}} \tag{III-23}$$

Then the first differential $dP[x,h]$ is evaluated to be

$$dP[x,h] = \frac{1}{p_1} \left(\int_0^T |y^d - A[x]|^{p_1} dt \right)^{\frac{1}{p_1}-1} p_1 \int_0^T |y^d - A[x]|^{p_1 - 1} \operatorname{sgn}(y^d - A[x]) \cdot$$

$$(-dA[x,h])\,dt + \lambda \frac{1}{p_2} \left(\int_0^T |x|^{p_2} dt \right)^{\frac{1}{p_2}-1} p_2 \int_0^T |x|^{p_2 - 1} \operatorname{sgn} x \; h \; dt \tag{III-24}$$

Defining

$$g(t) \triangleq |y^d - A[x]|^{p_1 - 1} \operatorname{sgn}(y^d - A[x]) \tag{III-25}$$

and introducing the adjoint differential operator $dA^*[x,g]$ defined by Equation (III-10), the gradient of $P[x]$ expressing the necessary condition for the existence of an extremum of the augmented performance criterion is written as:

$$\text{grad}P\lfloor x\rfloor = \lambda\left(\int_0^T |x|^{p_2} dt\right)^{\frac{1}{p_2}-1} |x|^{p_2-1}\text{sgn}x - \left(\int_0^T |y^d - A\lfloor x\rfloor|^{p_1} dt\right)^{\frac{1}{p_1}-1} dA^*\lfloor x,g\rfloor \tag{III-26}$$

The vanishing of the gradient of $P\lfloor x\rfloor$ leads to the relation

$$\lambda \frac{\|x\|_{p_2}}{\|x\|_{p_2}^{p_2}} |x|^{p_2-1}\text{sgn}x = \frac{\|y^d - A\lfloor x\rfloor\|_{p_1}}{\|y^d - A\lfloor x\rfloor\|_{p_1}^{p_1}} dA^*\lfloor x,g\rfloor \tag{III-27}$$

or

$$|x|^{p_2-1}\ \text{sgn}x = \frac{1}{\lambda} \frac{\|x\|_{p_2}^{p_2-1}}{\|y^d - A\lfloor x\rfloor\|_{p_1}^{p_1-1}} dA^*\lfloor x,g\rfloor \tag{III-28}$$

Solutions exist for the following cases of particular interest with regard to practical applications:

Case 1: Let $p_2 \to \infty$ and $p_1 \neq \infty$. Then from Equation (III-28)

$$|x| = \|x\|_{p_2}$$

$$\text{sgn}x = \text{sgn}(dA^*\lfloor x,g\rfloor) \tag{III-29}$$

and

$$x(t) = \max|x(t)|\ \text{sgn}(dA^*\lfloor x,g\rfloor) \qquad t \in (0,T) \tag{III-30}$$

since by definition

$$\|x\|_{p_2} = \max|x| \qquad \text{as } p_2 \to \infty \tag{III-31}$$

and λ , the Lagrange-multiplier, remains always positive; g(t) is interpreted according to Equation (III-25).

Case 2. Let $p_1 \to \infty$ and $p_2 = 2$. From Equation (III-28) again

$$x(t) = \frac{1}{\lambda} \|x\|_2 \, dA^* \lfloor x, g \rfloor \tag{III-32}$$

Here g(t) is given by Equation (III-25) and an appropriate interpretation of the value of g(t) is necessary as p_1 becomes a very large number. When the upper bound of the time interval (0,T) is unity the spaces corresponding to values of p_1 up to N, with N finite, are proper subspaces of $p_1 = \infty$. The same situation can be realized when $T \neq 1$ with a change of the independent variable t, as it is described in Section 1, Chapter II. Then a value for p_1 between five and ten will assure an excellent approximation to the desired expression for g(t).

Case 3. Finally with $p_1 = p_2 = 2$ (Hilbert Space) the solution for the control signal is given as

$$x(t) = \frac{1}{\lambda} \cdot \frac{\|x\|_2}{\|y^d - A\lfloor x \rfloor\|_2} \cdot dA^* \lfloor x, g \rfloor \tag{III-33}$$

with

$$g(t) = y^d - A\lfloor x \rfloor \tag{III-34}$$

A sequence of input functions can be constructed for each of Equations (III-30), (III-32) and (III-33) according to Equation (III-18) and the discussion presented above.

But, for the iterative solution of these nonlinear operator equations the adjoint differential $dA^*[x,g]$ must be identified and the Lagrange-multiplier λ evaluated. In the next two sections two techniques are described each leading to a solution of the identification problem.

III - 3. Identification of $dA^*[x,g]$, Biorthogonal Expansion Method.

Kulikowski [III-2] has suggested an expansion of the operator $dA^*[x,g]$ in terms of orthogonal time functions over the time interval (0,T) whenever all functions considered are elements of $L^2(0,T)$ (Hilbert) space. That is

$$dA^*[x,g] = \sum_i a_i \; h_i(t) \qquad t \in (0,T) \tag{III-35}$$

with

$$\frac{1}{T}\int_0^T h_i(t)h_j(t)\,dt = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{III-36}$$

and x,g and $h \in L^2(0,T)$.

Substituting the series representation (III-30) into the defining Equation (III-10) for the adjoint $dA^*[x,g]$ an expression for the unknown coefficients a_i in terms of plant differentials is obtained, i.e.

$$a_i = \frac{1}{T}\int_0^T g[x(t)]\, dA[x,h_i]\, dt \qquad i = 1, 2, \ldots \tag{III-37}$$

If the time functions considered are elements of a more general $L^p(0,T)$ - space then the adjoint differential $dA^*[x,g]$ may be expressed as the sum of an infinite series of biorthogonal time sequences, i.e.

$$dA^*[x,g] = \sum_i a_i \varphi_i(t) \tag{III-38}$$

with the sequences $\varphi_i(t)$ and $h_j(t)$ being biorthogonal if

$$(\ \varphi_i(t),\ h_j(t)\ ) = \delta_{ij} \tag{III-39}$$

and $h \in L^p(0,T)$, $\varphi \in L^q(0,T)$ and δ_{ij} is the Kronecker delta function. Substituting again the series representation (III-38) into the defining Equation for the adjoint differential $dA^*[x,g]$ and utilizing the biorthogonality property of the sequences $\varphi_i(t)$ and $h_j(t)$ we obtain the following expression for the unknown coefficients a_i in terms of plant output differentials:

$$a_i = \frac{1}{T}\int_0^T g[x(t)]\, dA[x,h_i]\, dt \qquad i = 1, 2, \ldots \tag{III-40}$$

One point should be clarified at this time: In an actual process operation the plant differential $dA[x,h]$ is approximated by

$$dA_\gamma[x,h] = \frac{1}{\gamma}\left\{A[x + \gamma h] - A[x]\right\} \tag{III-41}$$

where γ is small but finite, and of course, only a finite number m of terms in the expansion formula (III-38) can be used. Thus it becomes necessary to consider the convergence of $dA^*_{m,\gamma}[x,g]$, where

$$dA^*_{m,\gamma}[x,g] = \sum_{i=1}^{m} a_{\gamma i}\, \varphi_i(t) \tag{III-42}$$

and

$$a_{\gamma i} = \frac{1}{T}\int_0^T g(t)\, \frac{1}{\gamma}\left\{A[x+\gamma h_i] - A[x]\right\} dt \tag{III-43}$$

to the true value of the adjoint $dA^*[x,g]$. It will be shown that $dA^*_{m,\gamma}[x,g]$ converges uniformly to $dA^*[x,g]$ as m increases and $\gamma \to 0$, independent of x, i.e. for any function x(t) and given number ε, numbers M and G exist such that

$$\| dA^*_{m,\gamma}[x,g] - dA^*[x,g] \| \leq \varepsilon \tag{III-44}$$

when $m > M$ and $\gamma < G$. The following two lemmas are proved first:

Lemma 1. If $A[x]$ is a nonlinear operator, acting in the space of p^{th} integrable functions+, having a linear differintial $dA[x,h]$, the following relation holds:

$$\frac{1}{\gamma}\int_0^T g(t)\left\{ A[x+\gamma h] - A[x]\right\} dt = \int_0^T g(t)\,dA[x+\tau\gamma h, h]\,dt = \int_0^T h(t)\,dA^*[x+\tau\gamma h, g]\,dt \tag{III-45}$$

where τ is a number belonging to the interval (0,1).

Proof. Consider the function

$$\psi(\tau) = \int_0^T g(t) A[x+\tau\gamma h]\,dt \tag{III-46}$$

+A measurable function $x(t)$ defined on the interval $(0,T)$ belongs to $L^p(0,T)$-space, or otherwise is a function with integrable p^{th} power, if

$$\int_0^T |x(t)|^p\,dt < \infty$$

The integral is to be interpreted in the sense of Lebesque.

with real argument γ. The derivative of $\psi(\tau)$ is given by

$$\psi'(\tau) = \lim_{\Delta\tau\to 0} \frac{1}{\Delta\tau} \lfloor \psi(\tau+\Delta\tau) - \psi(\tau) \rfloor$$
$$= \int_0^T g(\ ) \lim_{\Delta\tau\to 0} \frac{1}{\Delta\tau} \left\{ A\lfloor x+\tau\gamma h+\Delta\tau\gamma h \rfloor - A\lfloor x+\gamma h \rfloor \right\} dt$$
$$= \int_0^T g(t)\ dA\lfloor x+\tau\gamma h, \gamma h \rfloor\ dt \qquad \text{(III-47)}$$

Applying the Lagrange formula to $\psi(\tau)$:

$$\psi(1) - \psi(0) = \psi'(\tau) \qquad 0 \le \tau \le 1 \qquad \text{(III-48)}$$

we obtain

$$\int_0^T g(t) \left\{ A\lfloor x+\gamma h \rfloor - A\lfloor x \rfloor \right\} dt = \int_0^T g(t) dA\lfloor x+\tau\gamma h, \gamma h \rfloor dt \qquad \text{(III-49)}$$

Since $dA\lfloor x,h \rfloor$ is a linear operator with respect to h:

$$\frac{1}{\gamma} \int_0^T g(t) \left\{ A\lfloor x+\gamma h \rfloor - A\lfloor x \rfloor \right\} dt = \int_0^T g(t) dA\lfloor x+\tau\gamma h, h \rfloor dt \qquad \text{(III-50)}$$

The second part of Equation (III-45) is obtained by direct substitution of the right-hand side of Equation (III-50) into the defining Equation for the adjoint differential.

For a more concise presentation of Lemma 2 and the main Theorem the following notation will be adopted:

Let

$$B\lfloor x \rfloor \triangleq dA^*\lfloor x,g \rfloor$$

and

$$dB\lfloor x,h \rfloor = d^2A^*\lfloor x,g,h \rfloor$$

It is assumed that the second plant differential $d^2A[x,h_1,h_2]$ exists for every x, h_1 and $h_2 \epsilon L^p$; then the second adjoint differential $d^2A^*[x,g,h]$ exists, i.e.

$$\left\| d^2A^*[x,g,h] \right\| < \infty \tag{III-51}$$

when $\|x\| < \infty$ and $\|h\| < \infty$, because, by definition

$$\int_0^T g(t)d^2A[x,h_1,h_2]dt = \int_0^T h_1(t)d^2A^*[x,g,h_2]dt \tag{III-52}$$

Lemma 2. If $B[x]$ is a nonlinear operator acting upon the p^{th} integrable space of functions and having the linear with respect to h(t) differential $dB[x,h] = B'[x]h$, where $B'[x]$ is a nonlinear derivative operator acting on h(t), then the following relation holds:

$$\frac{1}{|\gamma|} \left\| B[x+\gamma h]-B[x] \right\|_q \leq \sup_{0<\tau<1} \left\| B'[x+\tau\gamma h] \right\|_q \|h\|_p \tag{III-53}$$

Proof. If f(t) is an arbitrary function belonging to $L^p(0,T)$-space, from Lemma 1 it is true that

$$\frac{1}{\gamma}\int_0^T f(t)\left\{ B[x+\gamma h]-B[x] \right\} dt = \int_0^T f(t)dB[x+\tau\gamma h,h]dt \tag{III-54}$$

when $\tau \epsilon (0,1)$.

Applying Holder's inequality to both sides of Equation (III-54):

$$\left| \frac{1}{\gamma} \int_0^T f(t) \left\{ B[x+\gamma h]-B[x] \right\} dt \right| \leq \left| \frac{1}{\gamma} \right| \, \|f\|_p \, \|B[x+\gamma h]-B[x]\|_q \qquad \text{(III-55)}$$

and

$$\left| \int_0^T f(t) dB[x+\tau\gamma h,h] dt \right| \leq \|f\|_p \, \|dB[x+\tau\gamma h,h]\|_q \qquad \text{(III-56)}$$

The equality sign appears in Equation (III-56) when

$$f(t) = K \left| B[x+\gamma h]-B[x] \right|^{q-1} \operatorname{sgn}(B[x+\gamma h]-B[x]) \qquad \text{(III-57)}$$

Then from Equation (III-54) with f(t) as in (III-57) the following inequality holds true:

$$\left| \frac{1}{\gamma} \int_0^T f(t) \left\{ B[x+\gamma h]-B[x] \right\} dt \right| \leq \|f\|_p \, \|dB[x+\tau\gamma h,h]\|_q \qquad \text{(III-58)}$$

or

$$\begin{aligned} &\left| \frac{1}{\gamma} \int_0^T K \left| B[x+\gamma h]-B[x] \right|^{q-1} \left\{ B[x+\gamma h]-B[x] \right\} \operatorname{sgn}(B[x+\gamma h]-B[x]) dt \right. \\ &\qquad = \left| \frac{1}{\gamma} \int_0^T K \left| B[x+\gamma h]-B[x] \right|^{q} dt \right. \\ &\qquad \leq \left(\int_0^T \left| K \left| B[x+\gamma h]-B[x] \right|^{q-1} \operatorname{sgn}(B[x+\gamma h]-B[x]) \right|^{p} dt \right)^{\frac{1}{p}} \|dB[x+\tau\gamma h,h]\|_q \end{aligned}$$

$$\text{(III-59)}$$

And

$$\left|\frac{1}{\gamma}\int_0^T \left|B[x+\gamma h]-B[x]\right|^q dt\right| \leq \left(\int_0^T \left|B[x+\gamma h]-B[x]\right|^q dt\right)^{\frac{1}{p}} \left\|dB[x+\tau\gamma h, h]\right\|_q \tag{III-60}$$

Since $\frac{1}{p}+\frac{1}{q}=1$ the above inequality reduces to

$$\left|\frac{1}{\gamma}\int_0^T \left|B[x+\gamma h]-B[x]\right|^q dt\right| \leq \frac{\left\|dB[x+\tau\gamma h, h]\right\|_q}{\left\|B[x+\gamma h]-B[x]\right\|_q}\int_0^T \left|B[x+\gamma h]-B[x]\right|^q dt \tag{III-61}$$

and therefore

$$\left|\frac{1}{\gamma}\right| \left\|B[x+\gamma h]-B[x]\right\|_q \leq \left\|dB[x+\tau\gamma h, h]\right\|_q \tag{III-62}$$

with

$$dB[x+\tau\gamma h, h] = B'[x+\tau\gamma h]\, h \tag{III-63}$$

we may write (definition of the norm of an operator and Reference [II-1] p. 82):

$$\left|\frac{1}{\gamma}\right| \left\|B[x+\gamma h]-B[x]\right\|_q \leq \left\|B'[x+\tau\gamma h]\right\|_q \left\|h\right\|_p \tag{III-64}$$

or

$$\left|\frac{1}{\gamma}\right| \left\|B[x+\gamma h]-B[x]\right\|_q \leq \sup_{0<\tau<1} \left\|B'[x+\tau\gamma h]\right\|_q \left\|h\right\|_p \tag{III-65}$$

which proves the Lemma.

<u>Theorem</u>. If $A[x]$ is a nonlinear twice differentiable operator in L^p-space, then for any given function $x_k(t)$ of this space and any number ε there exist numbers M and G such that:

$$\left\| B_{m,\gamma}[x_k] - B[x_k] \right\| \leq \varepsilon \qquad \text{(III-66)}$$

where $m > M$, $\gamma < G$ and

$$B_{m,\gamma}[x_k] = \sum_{i=1}^{m} a_{\gamma i} \varphi_i(t) \qquad \text{(III-67)}$$

$$B[x_k] = \sum_{i=1}^{\infty} a_i \varphi_i(t) \qquad \text{(III-68)}$$

and

$$a_{\gamma i} = \frac{1}{\gamma T} \int_0^T g(t) \left\{ A[x_k + \gamma h_i] - A[x_k] \right\} dt \qquad \text{(III-69)}$$

$$a_i = \frac{1}{T} \int_0^T h_i(t) dA^*[x_k, g(t)] dt \qquad \text{(III-70)}$$

with $\varphi_i(t)$ and $h_i(t)$ being biorthogonal.

Proof. Using the Minkowski inequality[+] we may write:

$$\|B_{m,\gamma}[x_k]-B[x_k]\|_q \leq \|\sum_{i=1}^{m}[a_{\gamma i}-a_i]\varphi_i(t)\|_q + \|\sum_{i=m+1}^{\infty}a_i\varphi_i(t)\|_q \qquad \text{(III-71)}$$

Since the expansion of the q^{th} integrable function $B[x_k]$ into a series of $\varphi_i(t)$ functions converges, we can find corresponding to any given number $\varepsilon_1 < \varepsilon$ a number m such that

$$\|\sum_{i=m+1}^{\infty}a_i\varphi_i(t)\| \leq \varepsilon_1 \qquad \text{(III-72)}$$

From Lemma 1:

$$T[a_{\gamma i}-a_i] = \int_0^T h_i(t)\left\{B[x_k+\tau\gamma h_i]-B[x_k]\right\}dt \qquad \text{(III-73)}$$

and again making use of the Minkowski inequality together with the results of Lemma 2, we may write:

$$\begin{aligned} T\,\|\sum_{i=1}^{m}[a_{\gamma i}-a_i]\varphi_i(t)\|_q &\leq T\sum_{i=1}^{m}\|B[x_k+\tau\gamma h_i]-B[x_k]\|_q \\ &\leq \sum_{i=1}^{m}\|B'[x_k+\theta_i\tau\gamma h_i]\|_q\,\|\tau\gamma h_i\|_p\cdot T \\ &< Mm\,|\gamma|T \end{aligned} \qquad \text{(III-74)}$$

[+]See Reference [II-1] pp. 213-214.

where

$$M = \sup_{0 \le \theta_i < 1} \| B' [x_k + \theta_i \tau \gamma h_i] \|$$

Assuming that $\gamma < \frac{\varepsilon - \varepsilon_1}{M m T}$ we obtain

$$\| \sum_{i=1}^{m} [a_{\gamma i} - a_i] \varphi_i(t) \| < \varepsilon - \varepsilon_1 \qquad \text{(III-75)}$$

or

$$\| B_{m,\gamma} [x_k] - B[x_k] \| \le \varepsilon \qquad \text{(III-76)}$$

Q.E.D.

Practical considerations call for a proper choice of the biorthogonal time functions $\varphi_i(t)$. A method utilizing the fewest number of terms and closely approximating the operator $dA^*[x,q]$ makes use of an extension of the Gram-Schmidt orthogonalization process.

Let

$$f_{k-1}(t) = dA^*[x_{k-1}, g_{k-1}] \text{ and } f_k(t) = dA^*[x_k, g_k],$$

$k = 1, 2, \ldots$ If these two functions are similar in form and $f_{k-1}(t)$ has been identified in the past, the process for identification of $f_k(t)$ proceeds as follows:

If $h_1(t) = f_{k-1}(t)$, $h_2(t) = f'_{k-1}(t)$ ...

$h_m(t) = f_{k-1}^{(m-1)}(t)$, then

$$\varphi_1(t) = \frac{h_1(t)}{\|h_1(t)\|} \tag{III-77}$$

Take

$$h^*_2(t) = h_2(t) - u_1\, \varphi_1(t)$$

where the number u_1 is determined in such a way that $h_2^*(t)$ and $\varphi_1(t)$ are biorthogonal; application of the biorthogonality condition results in the following value for u_1:

$$u_1 = \frac{1}{T} \int_0^T h_2(t)\, \varphi_1(t) dt \tag{III-78}$$

and

$$\varphi_2(t) = \frac{h^*_2(t)}{\|h_2^*(t)\|} \tag{III-79}$$

Analogously the succeeding terms of the truncated series can be constructed. This procedure suffers from the disadvantage of not being self-starting. It is suggested therefore that $dA^*[x_1, g_1]$ be constructed from $\varphi_1(t)$ with the biorthogonal functions being arbitrarily chosen as sequences of unit vectors in the interval (0,T), or

$$\varphi_1(t) = (1, 0, 0, \ldots)$$
$$\varphi_2(t) = (0, 1, 0, 0, \ldots)$$
$$\varphi_3(t) = (0, 0, 1, 0, 0, \ldots) \quad \text{etc.}$$

For the second iteration of the overall adaptive process, i.e. for the construction of $dA^*[x_2, q_2]$, the function $f_1(t)$ is known and the Gram-Schmidt procedure can be utilized to advantage.

III - 4. Identification of $dA^*[x,q]$, Reverse Operator Method.

The biorthogonal expansion method is very general in scope and applies to all classes of systems that we are considering in the thesis. A disadvantage of this method is the length of time required for the evaluation of the coefficients of the truncated series. This disadvantage is overcome by the second method that we will discuss now. In one time interval from 0 to T we have an exact expression for the adjoint (in terms of the plant differential $dA[x,h]$). The price that we pay though is that this technique is applicable only to a small class of systems for which we require knowledge of the plant structure and the form of the nonlinearities. It is based on directly relating $dA^{*}\lfloor x,g \rfloor$ to the measurable plant differential $dA\lfloor x,h \rfloor$ through the use of a reverse operator R defined by:

$$R\left\{z(t)\right\} = z(T-t) \qquad t \in (0,T) \tag{III-80}$$

The operator R represents a device which stores a time function z(t) for $t \in (0,T)$ and emits it in reverse order during the succeeding T-time interval. Some properties of R are listed below:

$$\begin{aligned}
&1. \quad R\left\{R\left\{z(t)\right\}\right\} = z(t) \\
&2. \quad R\left\{az_1(t) + bz_2(t)\right\} = aR\left\{z_1(t)\right\} + bR\left\{z_2(t)\right\} \\
&3. \quad R\left\{u(t)\cdot v(t)\right\} = R\left\{u(t)\right\} R\left\{v(t)\right\} \\
&4. \quad R\left\{\int_0^t k(t-\tau)\, z(\tau)\, d\tau\right\} = \int_t^T k(\tau - t)\, R\left\{z(\tau)\right\} d\tau
\end{aligned} \tag{III-81}$$

The last property is established as follows:
From the definition, Equation (III-80), reverse operation means substitution of t by T - t. Then

$$R\left\{\int_0^t k(t-\tau)\, z(\tau)\, d\tau\right\} = \int_0^{T-t} k(T-t-\tau)\, z(\tau)\, d\tau \tag{III-82}$$

Letting $u = T-\tau$ the above expression can be written as

$$R\left\{\int_0^t k(t-\tau)\ z(\tau)\ d\tau\right\} = -\int_T^t k(-t+u)\ z(T-u)\ du \tag{III-83}$$

or

$$R\left\{\int_0^t k(t-\tau)\ z(\tau)\ d\tau\right\} = \int_t^T k(u-t)\ z(T-u)\ du \tag{III-84}$$

But

$$z(T-u) = R\left\{z(u)\right\}$$

which establishes property 4.

The identification method will now be illustrated assuming the following form for the plant operator:

$$A[x] = A_o + \int_0^t k(t-\tau)\ N[x(\tau)]\ d\tau \tag{III-85}$$

where $N[x(t)]$ represents a twice differentiable zero memory nonlinear device. The first differential of $A[x]$ is written as

$$dA[x,h] = \int_0^t k(t-\tau)\ \frac{dN}{dx}[x(\tau)]\ h(\tau)\ d\tau \tag{III-86}$$

Substituting Equation (III-86) into the defining equation for the adjoint operator:

$$\int_0^T g(t) \int_0^t k(t-\tau)\frac{dN}{dx}\lfloor x(\tau)\rfloor h(\tau)\, d\tau dt = \int_0^T h(t)\, dA^*\lfloor x,g\rfloor dt \qquad \text{(III-87)}$$

For physical system $k(t-\tau) = 0$ for $\tau > t$, Equation (III-87) is written as

$$\int_0^T g(t) \int_0^T k(t-\tau)\frac{dN}{dx}\lfloor x(\tau)\rfloor h(\tau) d\tau dt = \int_0^T h(t) dA^*\lfloor x,g\rfloor dt \qquad \text{(III-88)}$$

Interchanging the order of integration:

$$\int_0^T \frac{dN}{dx}\lfloor x(t)\rfloor h(t) \int_0^T k(\tau-t) g(t) d\tau dt = \int_0^T h(t) dA^*\lfloor x,g\rfloor dt \qquad \text{(III-89)}$$

But $k(\tau-t) = 0$ for $t > \tau$, therefore

$$\int_0^T \frac{dN}{dx}\lfloor x(t)\rfloor h(t) \int_t^T k(\tau-t) g(\tau) d\tau dt = \int_0^T h(t) dA^*\lfloor x,g\rfloor dt \qquad \text{(III-90)}$$

and

$$dA^*\lfloor x,g\rfloor = \frac{dN}{dx}\lfloor x(t)\rfloor \int_t^T k(\tau-t)\; g(\tau)\; d\tau \qquad \text{(III-91)}$$

Applying the reverse operator R to the expression for the plant differential $dA\lfloor x,h\rfloor$ given by Equation (III-86) results in

$$R\left\{dA\lfloor x,h\rfloor\right\} = \int_t^T k(\tau-t)\ R\left\{\frac{dN}{dx}\lfloor x(\tau)\rfloor\ h(\tau)\right\} d\tau \qquad \text{(III-92)}$$

Now introducing a specific variation $\bar{h}(t)$ given by :

$$\bar{h}(t) = \frac{R\left\{g(t)\right\}}{\frac{dN}{dx}\lfloor x(t)\rfloor} \qquad \text{(III-93)}$$

and by comparison with Equation (III-91) the adjoint differential operator $dA^{*}\lfloor x,g\rfloor$ is seen to be given by:

$$dA^{*}\lfloor x,g\rfloor = \frac{dN}{dx}\lfloor x(t)\rfloor\ R\left\{dA\lfloor x,\bar{h}\rfloor\right\} \qquad \text{(III-94)}$$

Similar expressions can be derived for other types of nonlinear plant operators. In general, for plants described by the sum of operators, delayed by T_o, of the form:

$$A\lfloor x\rfloor = \int_0^{t+T_o} k(t+T_o-\tau)\ N\lfloor x(T_o+\tau)\rfloor\ d\tau \qquad \text{(III-95)}$$

the adjoint differential operator can be obtained through the reverse operator method. The same procedure can be followed for nonlinear operators of the general form (III-3) which do not change when we substitute $\bar{t} = T-t$, $\bar{\tau}_i = T-\tau_i$ and interchange the order of integration. When in doubt whether a particular plant belongs to the class for which the adjoint operator can be determined by reversing in time the plant differential a test can be performed experimentally for every input x(t) using the following criterion:

$$\int_0^T h_1(t)\ dA\lfloor x,h_2\rfloor\ dt = \int_0^T h_2(t)\ R\left\{ dA\lfloor x,h_1\rfloor\right\} dt \qquad \text{(III-96)}$$

where $h_1(t)$ and $h_2(t)$ are arbitrary functions.

As it can be seen from Equation (III-94) the nonlinearity $N\lfloor x\rceil$ should be known within a constant vector of parameters which in turn is determined through steady-state plant measurements during the initial learning period $(0,t_1)$ as it is described by Pearson ⌊III-3⌉, or by one of the parameter identification techniques presented in Appendix I.

Having thus identified the adjoint differential operator, methods for solving nonlinear operator equations will be investigated before the total implementation scheme, which incorporates the evaluation of the Lagrange-multiplier λ, is presented.

The vanishing of the gradient of the performance functional, Equation (III-17), represents a nonlinear operator equation of the form

$$Q\lfloor x(t)\rceil = 0, \quad t \in (0,T) \qquad \text{(III-97)}$$

With an arbitrary initial element $x_o(t)$, $t \in (0,T)$, being specified, a sequence of inputs $x_n(t)$, $n = 1, 2, \ldots$ is sought, each element of the sequence being a function of time for $t \in (0,T)$, so that

$$P[x_{k+1}(t)] < P[x_k(t)] \qquad \text{(III-98)}$$

and the sequence converges strongly to the solution of Equation (III-97).

A number of techniques are available for the iterative solution of nonlinear operator equations. The convergence conditions for these techniques are extremely difficult to test beforehand for the type of operator equations arising in adaptive control problems. This is true even when the plant operator $A[x]$ is known explicitly.

Contraction mapping, Newton's methods and the technique introduced by Altman will be discussed in the next sections.

III - 5. Contraction Mapping.

For the iterative solution of Equation (III-97) using the principle of contraction mapping it is assumed that the operator equation under consideration can be rearranged in the form

$$x = B[x] \qquad \text{(III-99)}$$

where B is an operator mapping every $x \in X$ into the same space X. The mapping is said to be a contraction if $B[x]$ satisfies the Lipschitz condition

$$\rho(B[x_1] - B[x_2]) \leq \ell\rho(x_1, x_2) \qquad \text{(III-100)}$$

with $x_1, x_2 \in X$ and $\ell < 1$.

Then the solution to Equation (III-97) can be found by iteration:

$$x_n = B[x_{n-1}] \qquad n = 1, 2, \ldots \qquad \text{(III-101)}$$

according to the following Banach reducing transform principle [III-4].

<u>Theorem</u>. Let there be given in a complete metric space X an operator B which maps the elements of the space X into elements of the same space. Suppose further that (III-100) is satisfied, then there exists exactly one point x^o such that $B[x^o] = x^o$. The point x^o is called a fixed point of the operator B.

<u>Proof</u>. Starting with an arbitrary but fixed element $x_o \in X$, the sequence $\{x_n\}$ is constructed, putting

$$x_1 = B[x_o],\ x_2 = B[x_1],\ \ldots\ x_n = B[x_{n-1}]\ \ldots$$

It will be shown that the sequence $\{x_n\}$ is fundamental: It is true that

$$\rho(x_1,x_2) = \rho(B[x_o],B[x_1]) \leq b\rho(x_o,x_1) = b\rho(x_o,B[x_o])$$

$$\rho(x_2,x_3) = \rho(B[x_1,B[x_2]) \leq b\rho(x_1,x_2) \leq b^2\rho(x_o,B[x_o])$$

- -

$$\rho(x_n,x_{n+1}) \leq b^n\rho(x_o,B[x_o]$$

- -

Furthermore

$$\rho(x_n,x_{n+p}) \leq \rho(x_n,x_{n+1}) + \dots + \rho(x_{n+p-1},x_{n+p})$$

$$(b^n + b^{n+1} + \dots + b^{n+p-1}) \; \rho(x_o,B[x_o])$$

$$= \frac{b^n - b^{n+p}}{1-b}\rho(x_o,B[x_o])$$

According to the hypothesis $b < 1$, then

$$\rho(x_n,x_{n+p}) < \frac{b^n}{1-b}\rho(x_o,B[x_o])$$

As $n \to \infty$, $\rho(x_n,x_{n+p}) \to 0$, and therefore the sequence $\{x_n\}$ is fundamental. Since the space X is complete there exists an element $x^o \in X$ which is the limit of this sequence:

$$x^o = \lim_{n \to \infty} x_n$$

Now

$$\rho(x_n, B[x^o]) \leq \rho(x^o, x_n) + \rho(x_n, B[x^o]) = \rho(x^o, x_n) + \rho(B[x_{n-1}], B[x^o])$$

$$\leq \rho(x^o, x_n) + b\rho(x^o, x_{n-1})$$

If ε is arbitrarily given and n is sufficiently large, then

$$\rho(x^o, x_{n-1}) < \frac{\varepsilon}{2b} \quad \text{and} \quad \rho(x^o, x_n) < \frac{\varepsilon}{2}$$

Consequently

$$\rho(x^o, B[x^o]) < \varepsilon$$

But $\varepsilon > 0$ was arbitrary, hence this inequality holds if and only if

$$\rho(x^o, B[x^o]) = 0, \text{ i.e. if } B[x^o] = x^o.$$

If there are two elements $x^o \in X$, $y^o \in X$ for which

$$B[x^o] = x^o \quad \text{and} \quad B[y^o] = y^o$$

then

$$\rho(x^o, y^o) = \rho(B[x^o], B[y^o]) \leq b\rho(x^o, y^o)$$

Assuming that $\rho(x^o, y^o) > 0$ it follows that $\ell \geq 1$; this is however, a contradiction to the hypothesis. Thus uniqueness of the fixed point has been proved.

An estimate for the error of the n^{th} approximation in terms of the initial arbitrary element x_o and the Lipschitz constant ℓ is given by

$$\rho(x_n, x^o) \leq \frac{\ell^n}{1-\ell} \rho(x^o, B[x^o]) \tag{III-102}$$

From the results of Lemma 2, Section III-3, it is seen that the Lipschitz constant ℓ may be interpreted as the norm of the derivative operator of the operator B, i.e.

$$\ell = \sup_{x \in X} \| B'[x] \| \tag{III-103}$$

If the Lipschitz condition (III-100) is not satisfied the theorem does not deny the existence of a solution, it merely implies that the sequence $\{x_n\}$ defined by Equation (III-101) will not converge to a solution. For the special case when X is the real line and the operator B is an ordinary function the convergence of the sequence based on the contraction mapping principle can be easily verified geometrically (see Reference III-4, p. 44).

III - 6. Newton's Method.

Newton's method or the method of tangents is a well-known technique for the solution of real equations. Generalizations of Newton's method to the solution of operator equations in Banach spaces have been made by Kantorovich (Reference III-5, Chapter XVIII).

Kantorovich considered equations of the form given by (III-97) with Q an operator mapping an open set Ω of a Banach space X into another Banach space X. Q is assumed to have a zero in Ω , i.e. there exists an element x^o such that

$$Q[x^o] = 0 \quad \text{(III-104)}$$

Assuming further that the operator Q has a continuous derivative in Ω and taking any element $x_o \epsilon \Omega$, the element $Q[x_o] = Q[x_o] - Q[x^o]$ can be replaced by the approximation $Q'[x_o] (x_o-x^o)$, so that there is a basis for assuming that the solution of the equation

$$Q'[x_o] (x_o-x^o) = Q[x_o] \quad \text{(III-105)}$$

will be close to x^o. The solution of the linear Equation (III-105) is easily found to be

$$x_1 = x_o - (Q'[x_o])^{-1} Q[x_o] \quad \text{(III-106)}$$

substituting x_1 for x^o and on the obvious assumption that the inverse operation $(Q'[x_o])^{-1}$ exists.

Starting from the initial approximation x_o and continuing this process the sequence $\{x_n\}$ is obtained:

$$x_{n+1} = x_n - (Q'[x_n])^{-1} Q[x_n] \qquad n = 0, 1, 2, \ldots \tag{III-107}$$

Each x_n is an approximate solution of Equation (III-97) and in general becomes more accurate with increasing n.

Newton's method is not always practicable because x_n may pass outside the confines of the set Ω for a certain n, or the operation $(Q'[x_n])^{-1}$ may not exist. If the sequence $\{x_n\}$ converges to x^o and x_o is chosen sufficiently close to x^o, the operations $Q'[x_n]$ and $Q'[x_o]$ will differ only by a small amount in view of the continuity of Q'. Thus Equation (III-107) may be replaced by the simplified formula:

$$x'_{n+1} = x'_n - (Q'[x_o])^{-1} Q[x'_n] \qquad n = 0, 1, 2, \ldots; \quad x'_o = x_o \tag{III-108}$$

The process of forming the sequence $\{x'_n\}$ is called the modified Newton's method; it is much easier to use although in general yields worse approximations.

Kantorovich, in Reference III-5, gives a detailed treatment of the conditions under which Newton's method (original or modified) is practicable and convergent, i.e. the conditions under which the sequences $\{x_n\}$ and $\{x_n'\}$ converge to the solution of Equation (III-97).

III - 7. Altman's Method.

Due to the difficulty in obtaining the inverse of a linear operator in $L^p(0,T)$ space, as is required by Newton's method, M. Altman and P. Szeptycki [III-6,7] gave a different approach to the solution of nonlinear operator equations in L^p-spaces. The operator equation must be again of the form

$$Q[x] = 0 \tag{III-109}$$

with Q a nonlinear continuous operator defined on the sphere $S(x_o,r)$ with values in L^p. $S(x_o,r)$ is a closed sphere in L^p with center x_o and radius r.

Altman utilized the distinctive difference between operators and functionals in order to derive an iterative scheme equivalent to Newton's method but not requiring the determination of an inverse operator. Putting $F(x) = \|Q[x]\|^p$, Equation (III-109) reduces to

$$F(x) = 0 \tag{III-110}$$

Where F(x) is a functional. Assuming that $Q[x]$ is differentiable in the Frechet sense in the sphere $S(x_o,r)$ and applying a similar method as that of the previous section, an approximate process for the solution of Equation (III-109) is obtained as follows:

$$x_1 = x_o - \frac{\|Q[x_o]\|^p}{p\|f(x_o)\|^q} \cdot y_o$$
$$x_{n+1} = x_n - \frac{\|Q[x_n]\|^p}{p\|f(x_n)\|^q} \cdot y_n \qquad \text{(III-111)}$$

where

$f(x) = \bar{Q}'[x] \cdot |Q[x]|^{p-1} \operatorname{sgn} Q[x] \in L^q$

$Q'[x]$ is the Frechet differential of $Q[x]$

$\bar{Q}'[x]$ is the adjoint of $Q'[x]$, and

$y_n = |f(x_n)|^{\frac{1}{p-1}} \operatorname{sgn}[f(x_n)]$

Here, a set of conditions for choosing the y_n elements is sythesized first. It is possible to choose the y_n elements convenient to the particular application at hand and then analyze the convergence conditions. The simplest form for the application to the operator equations considered by Altman is constructed by choosing y_n as the image of the operator Q acting on x_n, i.e.

$$y_n = Q[x_n]$$

The iterations, in Hilbert space, take the form:

$$x_{n+1} = x_n - \frac{\| Q[x_n] \|^2}{2(Q'[x_n]\, Q[x_n],\, Q[x_n])} \cdot Q[x_n] \tag{III-112}$$

Altman, in a series of papers dealing with the solution of operator and functional equations, has presented various forms of the basic iterative algorithm and has analyzed the conditions under which the iterations converge to a solution. Of particular interest with regard to the class of problems under consideration in the thesis the following theorem is proved by Altman.

Theorem. The following conditions are to be satisfied:

1. The Frechet differential $Q'[x]$ of $Q[x]$ exists in the sphere $S(x_o, r)$ and satisfies condition

$$1/ \| f(x_o) \| \leq B_o \tag{III-113}$$

2. The Frechet differential $f'(x)$ exists and satisfies condition

$$\| f'(x) \| \leq k \quad \text{for every } x \text{ of } S(x_o, r) \tag{III-114}$$

3. The error estimate for the first approximate solution is known

$$\|x_1 - x_o\| = \frac{\|Q[x_o]\|^p}{p\|f(x_o)\|^q} \leq \eta_o \tag{III-115}$$

4. The constants B_o, η_o and k satisfy the inequality:

$$\frac{\|Q[x_o]\|^p}{p\|f(x_o)\|^q} \leq B_o \eta_o k = h_o \leq \frac{1}{2} \tag{III-116}$$

and

$$r = \frac{1-\sqrt{1-2h_o}}{h_o} \eta_o \tag{III-117}$$

Then Equation (III-109) has a solution x^o which belongs to the sphere $S(x_o,r)$ and the sequence of approximate solutions x_n defined by process (III-112) converges to x^o. An estimate of the error between x_n and the solution x^o is given by

$$\|x_n - x^o\| \leq \frac{1}{2^{n-1}} \cdot (2h_o)^{2^n-1} \tag{III-118}$$

The recurrence relation (III-111) can be expressed more directly, for each particular application, in terms of the criteria for optimal performance

Since $Q[x] = \nabla P[x]$ and

$$dP[x,h] = (h,Q[x])$$

The second differential $d^2P[x,h,h]$ may be computed as

$$d^2P[x,h,h] = (h,Q'[x]h)$$

Thus specifying particular variations of the input variable x(t) the quantities in Equation (III-111) may be expressed in terms of the first and second differentials of the performance functional which in turn can be evaluated from plant input-output measurements and Equations (III-6) and (III-7).

The techniques described above do not by any means exhaust the list of available methods for the iterative solution of functional and operator equations. Conjugate gradient techniques, Davidon's method and random search procedures in finite dimensional spaces have been extensively studied in recent years. Several accelerating algorithms have been applied in some cases.

Comparative studies at the present time tend to indicate strong dependence on the particular surface over which the search is performed, the dimensionality of the variables and the choice of an initial point. Second variational procedures like Newton's and Altman's methods are more accurate and converge faster than those techniques which require the evaluation of the gradient of the performance index only. They are, however, computationally more difficult to use and sensitive to the relative distance between the initial search point and the optimum. A discussion of and further comments on computational aspects, applicability and speed of convergence of some of these techniques are discussed in Chapter VI where applications of the adaptive control scheme are presented.

III - 8. The Implementation Scheme.

The solution proceeds in the following way:

1. During the initial time interval $(0,t_1)$ (Figure I-2) test inputs are applied to the plant and such items as plant type, nature of the nonlinearities present (whether even or odd) and the values of fixed gains and time constants necessary for the identification of the adjoint differential operator are determined. Results obtained during this learning period are also discussed in Chapter I.
2. The unconstrained problem is solved numerically, i.e. starting with an arbitrary initial input $x_o(t)$, $t \in (0,T)$, the next element $x_1(t)$ of the sequence $\{x_n\}$ is con-

structed from the condition of optimality ($\nabla P[x] = 0$) with $\lambda = 0$. Assuming its applicability in a particular application anyone of the iteration techniques described in the previous sections may be used for the explicit solution for the input $x_1(t)$. If x_1^* is the unconstrained solution then r_o is evaluated from

$$\| x_1^* \| = r_o \tag{III-119}$$

3. If $R \geq r_o$ (R is the value of the input constraint) then the solution to the constrained problem is given by the unconstrained solution.

4. If $R < r_o$ then operation on the boundary of the constraint will dictate that

$$\| x_1 \| = R \tag{III-120}$$

and the Lagrange-multiplier λ is evaluated from Equation (III-120). When operating on the boundary of the constraint the variation h(t) can be taken away from the boundary towards the constrained region so that inequality (III-5) is always satisfied.

When the input variation h(t) is fixed (as for example in the case of the identification of the adjoint differential through the reverse operator method) or the constraint itself is of the "hard" type allowing for no tolerance band outside the constraint surface then on-line application of the input test signal is not permitted and provision for a process model together with suitable switching instrumentation must be made. Means for obtaining the process model are described in Appendix I.

Steps two, three and four are repeated at each stage of the identification and optimization procedure to insure that the input constraint is satisfied throughout the operation of the system.

An analog model of the overall implementation scheme together with appropriate switching and logical decision functions is shown in Figure (III-3). The scheme employs the reverse operator method for the identification of the adjoint differential and proceeds under the assumption that the adjoint differential may be related to measurable plant output differentials and the known function g(t).

A specific process consisting of the cascade of a nonlinearity N and a linear part L with an input-output relation given by Equation (III-85) is utilized. For identification purposes the input signal is perturbed by a fixed variation as specified by Equation (III-93) and the adjoint differential is expressed by Equation (III-94). During the first T-time interval switch S_1 is in position 1, switch S_2 is in position 0 as is switch S_4 and switch S_3 is in position

1. With the period T assumed to be equal to or greater than the anticipated settling time of the process and the process itself being asymptotically stable the effect on the plant output of past inputs decays to zero. Thus, during the second T-time interval with S_1, S_3 and S_4 as before and switch S_2 in position 1, the process output depends only on x_o or equivalently the operator A is uniquely characterized by the process input during the interval (0,T) only. An appropriate delay line stores the output $A[x_o]$. The sequence of events during the third interval is as follows: S_1 in position 1, S_3 in position 2, while S_2 and S_4 are in their zero positions. The process output in the next T-time interval depends only on $x_o + \gamma \bar{h}$ and switch S_1 is still in position 1 with S_2 in position 2 and switch S_3 in position 2. The scheme now performs subtraction of $A[x_o]$ from $A[x_o + \gamma \bar{h}]$ and devision by γ giving a first order approximation to the output differential. Reversing in time the record of the output differential and multiplication of this by dN/dx results in an estimate for the adjoint differential $dA^*[x_o,g_o]$. Next the unconstrained problem ($\lambda = 0$) is solved numerically giving $x_1(t)$ in terms of the adjoint $dA^*[x_o,g_o]$. It is followed by testing for the input constraint condition; if $\|x_1\| \leq R$ switch S_4 is set to position 1 and S_1 to 2 and the constrained solution is the same as the unconstrained. If $\|x_1\| > R$ then the Lagrange-multiplier is evaluated and $x_1(t)$ is redeter-

mined in terms of dA^* and λ_1. Switch S_4 is set to position 2 in this case. Delay D_1 holds $x_1(t)$ for the next cycle of the implementation scheme. The identification part of the logical schematic is modified appropriately whenever the expansion method is used for the estimation of the adjoint.

III - 9. Multivariable Systems.

The extension of the adaptive optimal control scheme to plants consisting of a multiple of inputs and outputs can be made with relative ease. The formulation follows the procedure suggested by Pearson[III-8,9]. In [III-9] Pearson considers the general plant structure shown in Figure (III-4). H_1 and H_2 are vector valued operators generally nonlinear with memory. The identification problem for the plant structure of Figure (III-4) dictates a partitioning of the input vector into components and a knowledge of the manner each component affects the process output. That is, some of the inputs affect the outputs in a finite memory manner, some others feed directly into pure integrators which are preceded by only zero-memory odd functions and the last component feeds into pure integrators which are preceded by nonlinear operators possessing memory. For each class of input functions the identification problem is formulated in accordance with the outline presented in

Chapter I and this formulation has reciprocal effectson the optimization problem [III-10]. In this section a particular plant structure is considered,,shown in Figure (III-5), and consisting of n inputs and m outputs with $M_1 \ldots M_n$ and N being zero-memory nonlinear functions and $L_1 \ldots L_n$ linear operators of type zero. This assumption simplifies the presentation without any substantial loss of generality.

The i^{th} component of the output vector is related to the input vector through the nonlinear plant operator A according to the relationship:

$$y_i(t) = A_i[x_1(t),\ x_2(t), \ldots\ x_n(t)] \qquad t \in (0,T), \quad i = 1, 2, \ldots m \tag{III-121}$$

or, in a more compact form:

$$\underline{y}(t) = \underline{A}[\underline{x}] \tag{III-123}$$

The vector valued operator $\underline{A}$ is assumed to have a weak differential $d\underline{A}[\underline{x},\ \underline{h}]$, given by:

$$d\underline{A}[\underline{x},\ \underline{h}] = \frac{\partial}{\partial \gamma} A[\underline{x} + \gamma \underline{h}]\Big|_{\gamma = 0} \tag{III-123}$$

Furthermore the differential $d\underline{A}[\underline{x},\ \underline{h}]$ is assumed to be linear with respect to the variations $\underline{h}(t)$ and may be considered as an m x n matrix of derivative operators operating on the vector $\underline{h}$, or:

$$d\underline{A}\lfloor\underline{x},\underline{h}\rfloor = A'\lfloor x\rfloor\cdot\underline{h} = \begin{bmatrix} A'_{11} & A'_{12} & \cdots & A'_{1n} \\ \vdots & \vdots & & \vdots \\ A'_{m1} & A'_{m2} & \cdots & A'_{mn} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \qquad \text{(III-124)}$$

where $A'_{ij} = A'_{ij}\lfloor\underline{x}\rfloor$ represents the derivative operator for $A_i\lfloor\underline{x}\rfloor$ corresponding to a variation in x_j <u>only</u>, i.e.

$$A'_{ij}\lfloor\underline{x}\rfloor\, h_j = \lim_{\gamma\to 0} \frac{A_i\lfloor x_1,\, x_2, \ldots\, x_j + \gamma h_j \ldots\, x_n\rfloor - A_i\lfloor\underline{x}\rfloor}{\gamma} \qquad \text{(III-125)}$$

The i^{th} differential of the vector operator $\underline{A}\lfloor\underline{x}\rfloor$ follows from Equation (III-124) as:

$$dA_i\lfloor\underline{x},\, \underline{h}\rfloor = \sum_{j=1}^{n} A'_{ij}\lfloor\underline{x}\rfloor\cdot h_j \qquad \text{(III-126)}$$

$$i = 1,\, 2,\, \ldots\, m$$

The performance functional for the multidimensional system, in correspondence with Equation (III-4) for the single input-output case, is given by:

$$J\lfloor\underline{x}\rfloor = \int_0^T G(\underline{x},\, \underline{y},\, \underline{y}^d)\, dt \qquad \text{(III-127)}$$

The differential of Equation (III-127) is written as:

$$dJ[\underline{x},\ \underline{h}] = \int_0^T \left\{ \frac{\partial G}{\partial \underline{x}} \underline{h} + \frac{\partial G}{\partial \underline{y}} d\underline{A}[\underline{x},\ \underline{h}] \right\} dt \tag{III-128}$$

where $\frac{\partial G}{\partial \underline{x}}$ is an n-dimensional vector and $\frac{\partial G}{\partial \underline{y}}$ is an m-dimensional vector with components respectively:

$$\left\{ \frac{\partial G}{\partial \underline{x}} \right\}_i = \frac{\partial}{\partial x_i} G(\underline{x}, \underline{y}, \underline{y}^d)$$

$$\left\{ \frac{\partial G}{\partial \underline{y}} \right\}_j = \frac{\partial}{\partial y_j} G(\underline{x}, \underline{y}, \underline{y}^d) \qquad \begin{matrix} i = 1,\ 2, \ldots n \\ j = 1,\ 2, \ldots m \end{matrix} \tag{III-129}$$

An n x m matrix of adjoint derivative operators $A'^*[\underline{x}]$ will be defined via the following inner product relationship so that the vector of arbitrary variations $\underline{h}(t)$ does not appear in the expression for the necessary condition of optimality.

$$\int_0^T \underline{g}(t)\ A'[\underline{x}]\ \underline{h}\ dt = \int_0^T \underline{h}(t)\ A'^*[\underline{x}]\ \underline{g}\ dt \tag{III-130}$$

The elements of $A'^*[\underline{x}]$ are recognized to be the adjoints of the corresponding elements of the matrix $A'[\underline{x}]$ transposed. In matrix form, designating the adjoint of A'_{ij} by A'^*_{ij} :

$$A'^{*}[\underline{x}]\ \underline{g}(t) = \begin{bmatrix} A'^{*}_{11} & A'^{*}_{21} & \cdots & A'^{*}_{m1} \\ \vdots & \vdots & & \vdots \\ A'^{*}_{1n} & A'^{*}_{2n} & \cdots & A'^{*}_{mn} \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix} \tag{III-131}$$

Substitution of Equation (III-130) into Equation (III-128) results in:

$$dJ[\underline{x},\underline{h}] = \int_0^T \underline{h}(t)\left\{ \frac{\partial G}{\partial \underline{x}} + A'^{*}[\underline{x}] \cdot \frac{\partial G}{\partial \underline{y}} \right\} dt \tag{III-132}$$

Since $\underline{h}(t)$ is an arbitrary vector function in the interval (0,T) the necessary condition for the existence of an extremum of $J[\underline{x}]$ is written as:

$$\frac{\partial G}{\partial \underline{x}} + A'^{*}[\underline{x}] \frac{\partial G}{\partial \underline{y}} = 0 \tag{III-133}$$

Each component of the n-dimensional vector gradient expressed by Equation (III-133) must vanish in the interval (0,T).

Techniques for the identification of the matrix of adjoint differential operators and for the iterative solution of the vector operator Equation (III-133) are similar to those discussed in previous sections for the one-dimensional case. As it was pointed out already, knowledge of the manner in which the input functions will affect the outputs is necessary; then each input is perturbed separately and the corresponding elements of the adjoint matrix are identified. A multi-dimensional search technique is required for the solution of Equation (III-133).

The formulation, solution and implementation of the conditional optimization problem follows directly from the single-input, single-output case and the discussion above. It is only necessary to define an appropriate vector norm in L^p-spaces. If $\underline{x}(t) \in X$ (X an L^p-space) is an n-dimensional vector then

$$\|\underline{x}\|_p = \left(\int_0^T \sum_{i=1}^{n} |x_i(t)|^p \, dt \right)^{\frac{1}{p}} \qquad p \geq 1 \qquad \text{(III-134)}$$

An analogous norm is defined for the output vector $\underline{y}(t)$. The performance criterion may be defined accordingly and the formulation and solution utilize the concepts put forth above. The detailed analysis is presented in the next chapter in connection with the adaptive optimal control problem for plants described by nonlinear differential equations.

FIGURE III-1

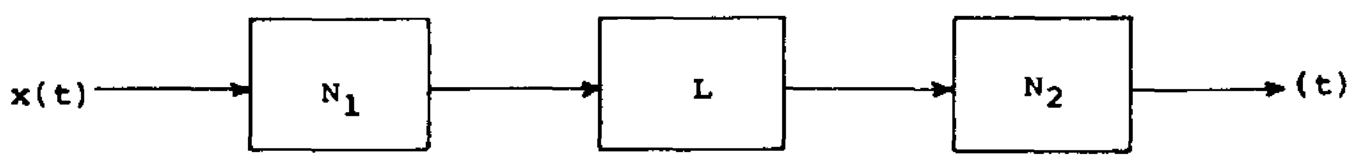


FIGURE III-2

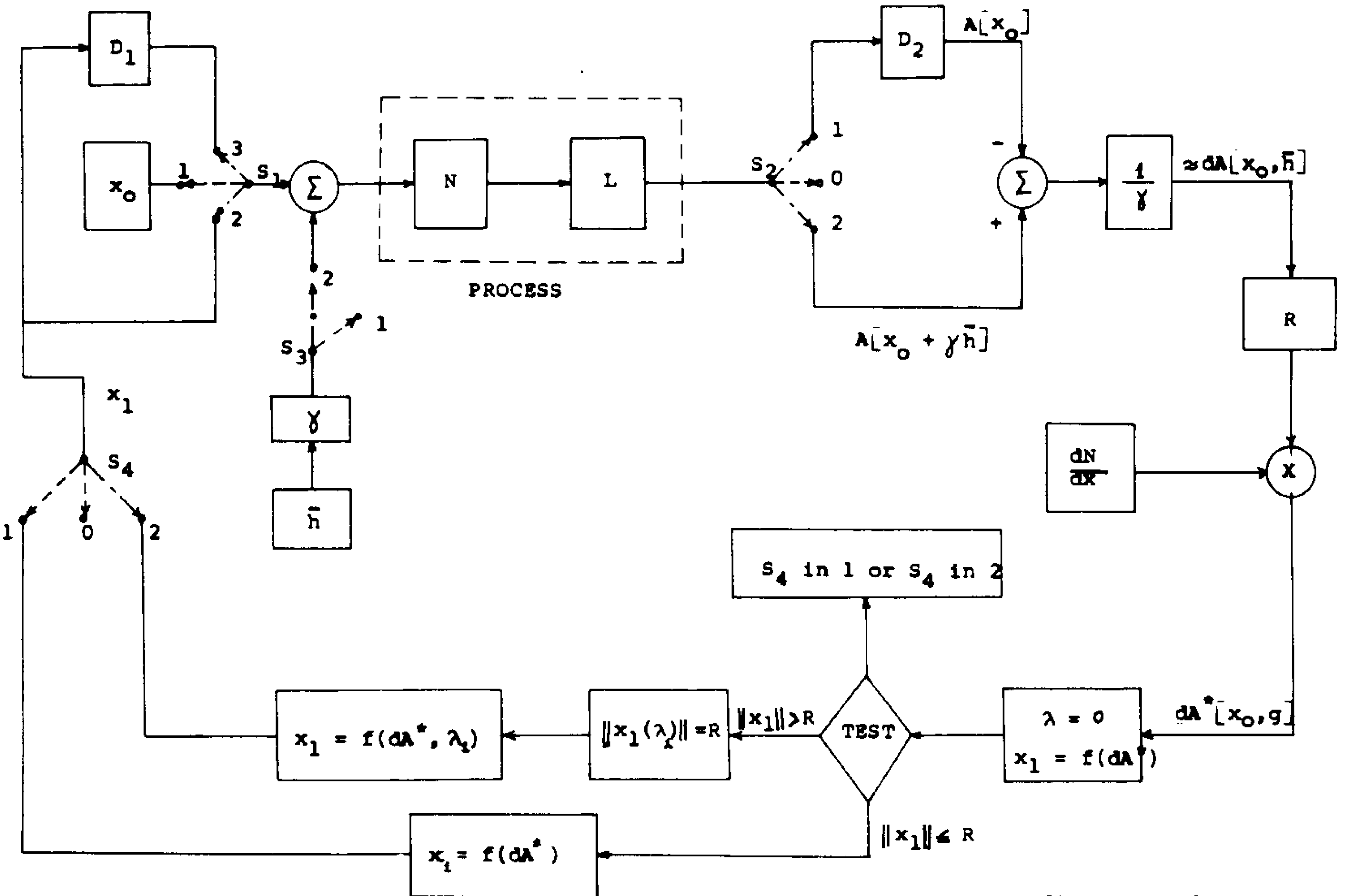

D1
x0
1
3
2
S1
Σ
N
L
PROCESS
2
1
S3
γ
h̄
S2
1
0
2
D2
A[x0]
-
+
Σ
A[x0 + γh̄]
1/γ
≈dA[x0,h̄]
R
x1
S4
1
0
2
dN/dX
X
S4 in 1 or S4 in 2
TEST
‖x1‖>R
‖x1(λi)‖=R
x1 = f(dA*, λi)
λ = 0
x1 = f(dA*)
dA*[x0,g]
‖x1‖≤ R
xi = f(dA*)

FIGURE III-3

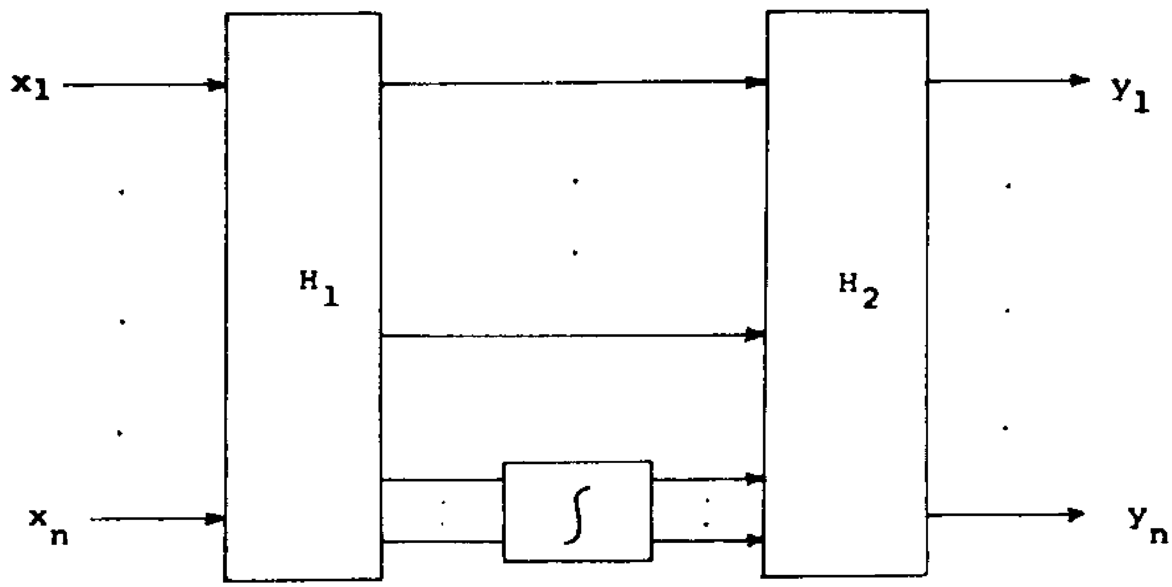


FIGURE III-4

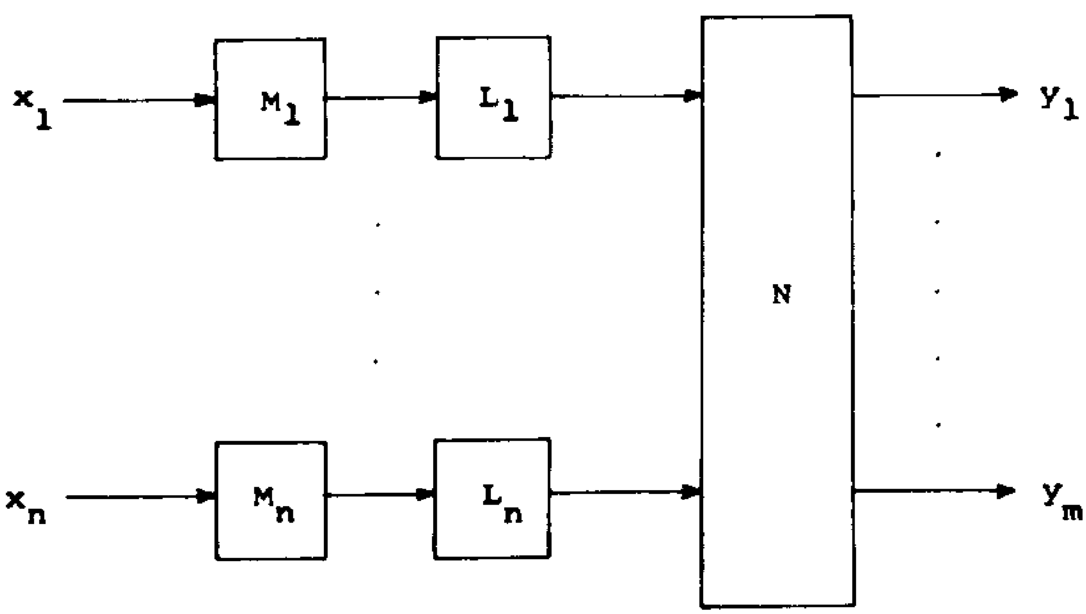


FIGURE III-5

BIBLIOGRAPHY

III - 1. R. Kulikowski, " On Optimization of Nonlinear Dynamic Processes with Unknown Parameters ", Technical Report No. 71, Columbia University, School of Engineering, 1962.

III - 2. R. Kulikowski, " On Adaptive Optimization of Nonlinear Random Control Systems ", (unpublished report).

III - 3. A.E. Pearson, " Adaptive Optimal Control of Nonlinear Processes ", Proceedings of the JACC, 1963.

III - 4. L.A. Liusternik and V.I. Sobolev, " Elements of Functional Analysis ", (book) Frederick Ungar Publishing Company, New York, 1961.

III - 5. L.V. Kantorovich and G.P. Akilov, "Functional Analysis in Normed Spaces", The MacMillan Co., New York, 1964.

III - 6. M. Altman, " On the Approximate Solution of Operator Equations in L^p-spaces ", Bulletin De L' Academie Polonaise Des Sciences, cl. III, vol. V, No. 12, 1957, pp. 1099 - 1103.

III - 7. P. Szeptycki, " A Remark on the Method of M. Altman of Solving Nonlinear Equations in L^p-Space ", <u>Bulletin De L' Academie Polonaise Des Sciences</u>, cl. III, vol. V, No. 12, pp. 1109 - 1112.

III - 8. A.E. Pearson, " Adaptive Optimal Control of Nonlinear Processes ", Ph. D. Dissertation, Columbia University, 1963.

III - 9. A.E. Pearson, " Adaptive Optimal Steady State Control of Nonlinear Systems ", in <u>Advances in Control Systems</u>, vol. 5, 1967.

III - 10. A.E. Pearson and P.E. sarachik, " On the Formulation of Adaptive Optimal Control Problems ", Proc. of JACC, pp. 13 - 22, 1964.

CHAPTER IV

DIFFERENTIAL EQUATION FORMULATION

In the previous chapter an integral operator relates the output to the input function. But realistic mathematical models of control processes, in most instances, are given in terms of a set of nonlinear differential equations. In this chapter it is assumed that the plant is described by such a set of equations. The differential equations are assumed to contain a vector of slowly - varying unknown parameters.

The optimization problem is again formulated using classical calculus of variations techniques and the solution is obtained in the form of an iteration. For the measurement of plant differentials the procedure suggested by Pearson and Sarachik [IV-1] and utilized in the integral equation formulation is followed. The identification of the adjoint plant differentials (necessary for the construction of the sequence of input signals) is accomplished using both the biorthogonal expansion method and the properties of the reverse operator.

A wide range of performance criteria is again feasible through the use of the concepts of nonlinear functional analysis, and the case of bounded input signals is included. Finally, the analysis here is directly carried out for plants characterized by a vector input - output relationship.

IV - 1. Statement of the Problem.

It is assumed that the plant to be controlled is described by:

$$\dot{\underline{y}}(t) = \underline{f}(\underline{x}, \underline{y}, \underline{c}) \tag{IV-1}$$

where

$\underline{x}(t)$ is an n-dimensional input vector

$\underline{y}(t)$ is an m-dimensional output vector

$\underline{c}(t)$ is a vector of slowly varying unknown parameters of dimensionality k, and $\underline{f}$ is differentiable with respect to both $\underline{x}$ and $\underline{y}$ and these differentials are bounded and continuous in a finite time interval (0,T). These conditions will assure the existence of an explicit solution for the control signal x(t) in the interval (0,T).

A performance index is given of the form:

$$J(\underline{x}) = \int_0^T G(\underline{x}, \underline{y}, \underline{y}^d)dt \tag{IV-2}$$

where $\underline{y}^d$ is a fixed vector of desired output values (same or lower dimensionality than the output vector $\underline{y}(t)$) and G possesses continuous and bounded differentials with respect to all of its arguments $\underline{x}$ and $\underline{y}$. $J(\underline{x})$ expresses the quality of system performance in the interval (0,T).

The objective is to find a sequence of input signals $\{\underline{x}^*(t)\}$, $t \in (0,T)$, so that the limiting element of this sequence minimizes (maximizes) the criterion function $J(\underline{x})$. The latter, being a measure of the overall plant performance, may take different forms depending on specific plant requirements. For example, in regulator problems

$$J(\underline{x}) = \| \underline{y}^d - \underline{y}(t) \|_p \qquad \text{(IV-3)}$$

where $\|z\|$ is the norm of z in L^p-space, and a measure of the difference between the desired and the actual output values is minimized. In some applications the total output product is to be maximized.
Then:

$$J(\underline{x}) = \| \underline{y}(t) \|_p \qquad \text{(with } p = 1) \qquad \text{(IV-4)}$$

Additionally the input signals will be constrained in magnitude to be less than a constant value R, or

$$\| \underline{x}(t) \|_p \leq R \qquad (\text{with } p \to \infty) \qquad \text{(IV-5)}$$

Other forms for the performance index and the input constraint are possible as it is shown in Chapter III. The particular expressions of Equations IV-3, 4 and 5 were chosen here because they lead to simple and meaningful solutions, from a practical point of view. The practical implications of this kind of formulation are investigated in Chapter VI.

With the inequality constraint (IV-5) the problem is reformulated as follows: Find that value $\underline{x}(t)$ which maximizes (or minimizes) $J(\underline{x})$ while still satisfying (IV-5).

IV - 2. Method of Solution.

It is supposed that a sufficient criterion of performance is the maximization of the plant output product, i.e. Equation (IV-4) is utilized. (The analysis proceeds along similar lines for other performance criteria). Employing the method of Lagrange multipliers, the right - hand side of Equation (IV-4) is augmented and the following synthetic functional is considered:

$$P[\underline{x}] = \|\underline{y}(t)\|_{p_1} + \lambda \|\underline{x}(t)\|_{p_2} \qquad \text{(IV-6)}$$

where $\lambda > 0$ is the Lagrange-multiplier and $p_1 = 1$, $p_2 \to \infty$.

The vector norms are given explicitly by:

$$\|\underline{y}(t)\|_{p_1} = \left(\int_0^T \sum_{i=1}^{m} |y_i(t)|^{p_1} dt\right)^{\frac{1}{p_1}} \qquad \text{(IV-7)}$$

and

$$\left\| \underline{x}(t) \right\|_{p_2} = \left(\int_0^T \sum_{i=1}^{n} \left| x_i(t) \right|^{p_2} dt \right)^{\frac{1}{p_2}} \qquad \text{(IV-8)}$$

Sarachik and Kranc [IV-2] introduced a more general norm which enables each component of $\underline{y}$ and $\underline{x}$ to be constrained separately, but in the present case the more simple formulation of (IV-7) and (IV-8) suffices.

The necessary condition for the existence of an extremum of the performance functional (IV-6) is the vanishing of its first differential. The sufficiency condition dictates that the second differential $d^2P[\underline{x}, \underline{h}, \underline{h}]$ be strictly negative for a relative maximal value of the quality functional (Chapter II, Section 6). For functionals of several variables it is difficult to state such a condition explicitly, and of course much harder to prove and/or implement it. It will be considered here as being satisfied; moreover the problems are again physically motivated and it should be intuitively clear whether the performance criterion is minimized or maximized. Finally, the value of the performance functional may be checked at each stage of the optimization process thus settling the question of the nature of the extremum.

The first differential of $P[\underline{x}]$ is given by:

$$dP[\underline{x},\underline{h}] = \left(\int_0^T \sum_{i=1}^{m} |y_i(t)|^{p_1} dt \right)^{\frac{1}{p_1}-1} \int_0^T \underline{g}(t) dy[\underline{x}]\, \underline{h}^T(t) dt$$

$$+ \left(\int_0^T \sum_{i=1}^{n} |x_i(t)|^{p_2} dt \right)^{\frac{1}{p_2}-1} \int_0^T \underline{h}(t)\, \underline{\ell}[\underline{x}] dt \qquad \text{(IV-9)}$$

where $\underline{h}(t)$ is an n-dimensional row vector whose elements are arbitrary variations belonging to the input space X. And by definition:

$$\underline{\ell}[\underline{x}] = \left(|x_1|^{p_2-1} \operatorname{sgn} x_1,\ |x_2|^{p_2-1} \operatorname{sgn} x_2, \ldots |x_n|^{p_2-1} \operatorname{sgn} x_n \right)^T \qquad \text{(IV-10)}$$

$$\underline{g}(t) = \left(|y_1|^{p_1-1} \operatorname{sgn} y_1,\ |y_2|^{p_1-1} \operatorname{sgn} y_2, \ldots |y_m|^{p_1-1} \operatorname{sgn} y_m \right) \qquad \text{(IV-11)}$$

The symbol $(\ )^T$ denotes transposition of the rows and columns of a vector or matrix. Also:

$$dy[\underline{x}]\, \underline{h}^T(t) = \begin{bmatrix} dy_{11}[\underline{x}] & dy_{12}[\underline{x}] & \cdots & dy_{n}[\underline{x}] \\ \vdots & \vdots & & \vdots \\ dy_{m1}[\underline{x}] & dy_{m2}[\underline{x}] & \cdots & dy_{mn}[\underline{x}] \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \qquad \text{(IV-12)}$$

In general $dy_{ij}[\underline{x}]$ represents the derivative operator for $y_i[\underline{x}]$ corresponding to a variation in x_j only, i.e.

$$dy_{ij}[\underline{x}] = \lim_{\gamma \to 0} \frac{y_i[x_1, x_2 \ldots x_j + \gamma h_j \ldots x_n] - y_i[\underline{x}]}{\gamma} \qquad \text{(IV-13)}$$

Introducing the n x m matrix of adjoint derivative operators $dy^*[\underline{x}]$ through the inner product form:

$$\int_0^T \underline{g}(t)\; dy[\underline{x}]\; \underline{h}^T(t)\; dt = \int_0^T \underline{h}(t)\; dy^*[\underline{x}]\; \underline{g}^T(t)\; dt \qquad \text{(IV-14)}$$

with

$$dy^*[\underline{x}]\; \underline{g}^T(t) = \begin{bmatrix} dy^*_{11}[\underline{x}] & dy^*_{12}[\underline{x}] & \ldots & dy^*_{1m}[\underline{x}] \\ & & & \\ dy^*_{n1}[\underline{x}] & dy^*_{n2}[\underline{x}] & \ldots & dy^*_{nm}[\underline{x}] \end{bmatrix} \begin{bmatrix} g_1 \\ \\ g_m \end{bmatrix} \qquad \text{(IV-15)}$$

and defining the gradient of the performance functional by:

$$dP[\underline{x}, \underline{h}] = \int_0^T \underline{h}(t)\; \underline{\text{grad}}P[\underline{x}]dt \qquad \text{(IV-16)}$$

since $\underline{h}(t)$ is arbitrary, the vanishing of the gradient is an equivalent necessary condition for the existence of an extremum of $P[\underline{x}]$.

For $p_1 = 1$ and by substitution of Equation (IV-14) into the expression for the first differential of $P[\underline{x}]$, Equation (IV-9), and finally through utilization of Equation (IV-16) the vanishing of the gradient of $P\lfloor \underline{x} \rfloor$ is directly expressed by:

$$\underline{\text{grad}}P\lfloor \underline{x} \rfloor = dy^*\lfloor \underline{x} \rfloor \, \underline{g}^T(t) + \lambda\left(\int_0^T \sum_{i=1}^{n} \left| x_i(t) \right|^{p_2} dt\right)^{\frac{1}{p_2}-1} \underline{\ell}\lfloor \underline{x} \rfloor = 0 \qquad \text{(IV-17)}$$

Each element of the vector $\text{grad}P\lfloor \underline{x} \rfloor$ must vanish in the interval (0,T). A sequence of input functions can now be constructed with each member of the sequence computed from the previous one according to the relationship:

$$\underline{x}_{k+1} = S(\underline{x}_k,\ dy^*\lfloor \underline{x}_k \rfloor) \qquad \text{(IV-18)}$$

The limiting sequence results in the optimal input vector $\underline{x}^*(t)$, $t \in (0,T)$, which satisfies (IV-17).

For the iterative solution of the vector nonlinear operator equation (IV-17) the matrix of adjoint differential operators $dy^*\lfloor \underline{x} \rfloor$ must be identified and the Lagrange-multiplier λ evaluated. Before methods for evaluation of λ and the implementation of the adaptive scheme are presented, two techniques will be described in the following sections leading to a solution of the identification problem.

IV - 3. Identification of the Adjoint Differential Operator.

The adjoint differential operator will be identified through direct monitoring of the plant output vector and measurement of first order plant differentials. Specifically, the output vector observed for a period of time equal to T with a known vector $\underline{x}(t)$ applied to the plant input. During a subsequent time interval (0,T) the i^{th} element of the input vector is perturbed by the variation $\delta x_i = \gamma h_i$ and a measure of the variation of the j^{th} output component is obtained by:

$$\delta y_{ji}(\underline{x}, \delta x_i) = y_j(x_1, x_2, \ldots x_i + \delta x_i, \ldots x_n) - y_j(\underline{x}) \qquad \text{(IV-19)}$$

With $\delta x_i = \gamma h_i$ and division of the expression on the right-hand side of Equation (IV-19) by γ, while keeping γ small, yields a first order approximation to the differential $dy_{ji}(\underline{x}, h_i)$. The procedure just described is applicable only to a small class of systems, i.e. type zero, with no accumulation present in the plant and only after the latter has reached a "steady state" with respect to the input $\underline{x}(t)$. Pearson and Sarachik [IV-1] have presented methods for the identification of the differential operator when the plant is of type other than zero; the time required for identification increases with increasing number of pure accumulators present in the plant structure as it is described in Chapter I. For simplicity it is assumed here that the plant is of type zero.

Substitution of the variational form for both input and output variables into Equation (IV-1) and retainment of the first order terms of the Taylor's series expansion results in:

$$\frac{d}{dt} \underline{\delta y}_{mi} = F_i(t)\, \underline{\delta y}_{mi} + H_i(t)\, \underline{\delta x}_i \qquad \text{(IV-20)}$$

for $i = 1, 2, \ldots n$

where $\underline{\delta y}_{mi}$ is an m-dimensional column vector and $F_i(t)$, $H_i(t)$ are generally time-varying matrix multipliers in the interval (0,T) functions of the parameter vector $\underline{c}$ and given by:

$$F_i(t) = \begin{bmatrix} \dfrac{\partial f_1[\underline{y},\underline{x},\underline{c}]}{\partial y_1} & \cdots & \dfrac{\partial f_1[\underline{y},\underline{x},\underline{c}]}{\partial y_m} \\ \\ \dfrac{\partial f_m[\underline{y},\underline{x},\underline{c}]}{\partial y_1} & & \dfrac{\partial f_m[\underline{y},\underline{x},\underline{c}]}{\partial y_m} \end{bmatrix} \qquad \text{(IV-21)}$$

and

$$H_i(t) = \begin{bmatrix} \frac{\partial f_1[\underline{y},\underline{x},\underline{c}]}{\partial x_i} & 0 & \cdot & \cdot & 0 \\ 0 & \frac{\partial f_2[\underline{y},\underline{x},\underline{c}]}{\partial x_i} & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & & & & \frac{\partial f_m[\underline{y},\underline{x},\underline{c}]}{\partial x_i} \end{bmatrix} \qquad \text{(IV-22)}$$

If the parameter vector $\underline{c}$ is known, $F_i(t)$ and $H_i(t)$ may be computed from known values of the input and output vector functions. The linearized Equation (IV-20) is one of a series of vector equations which are obtained as i takes on values from 1 to n, i.e. as each input is separately excited.

If $\Phi(t,\tau)$ is the transition matrix corresponding to Equation (IV-20) and matrices $F_i(t)$ and $H_i(t)$ are assumed to be bounded and continuous in a closed finite interval [IV-3], then the solution to the system of linear differential equations (IV-20) may be expressed in the following form:

$$\underline{\delta y}_{mi}(t) = \Phi(t,0)\,\underline{\delta y}_{mi}(0) + \int_0^t \Phi(t,\tau) H_i(\tau)\,\underline{\delta x}_i(\tau)\,d\tau$$

$$\text{for } i = 1, 2, \ldots n \qquad \text{(IV-23)}$$

or, in a more compact form, as:

$$\underline{\delta y}_{mi}(t) = A\,\underline{\delta y}_{mi}(0) + B\,\underline{\delta x}_i \qquad \text{(IV-24)}$$

with A and B representing linear operators. The inner product defined by the left-hand side of Equation (IV-14) will be designated by:

$$\langle \underline{g}(\underline{y}),\ \underline{\delta y}_{mi}(t)\rangle = \langle \underline{g}(\underline{y}),\ A\,\underline{\delta y}_{mi}(0) + B\,\underline{\delta x}_i \rangle \qquad \text{(IV-25)}$$

Utilizing property 2, Equation (II-19), of the inner product form, Equation (IV-25) is written as:

$$\langle \underline{g}(\underline{y}),\ \underline{\delta y}_{mi}(t)\rangle = \langle \underline{g}(\underline{y}),\ A\,\underline{\delta y}_{mi}(0)\rangle + \langle \underline{g}(\underline{y}),\ B\,\underline{\delta x}_i\rangle \qquad \text{(IV-26)}$$

To identify the differentials of the plant output it is necessary to drive the plant into the "steady state" condition (for example by applying a sequence of identical inputs for successive T-time intervals) resulting in identical initial conditions for the same T-time intervals. Thus the first term of the right-hand side of Equation (IV-26) will vanish.

By definition:

$$\langle \underline{g}(\underline{y}),\ B\,\underline{\delta x}_i\rangle = \langle \underline{\delta x}_i^T,\ B^*\underline{g}(\underline{y}) \rangle \qquad \text{(IV-27)}$$

With $B\,\underline{\delta x_i}$, in general, given by:

$$B\,\underline{\delta x_i} = \int_{t_0}^{t} \Phi(t,\tau)H_i(\tau)\,\underline{\delta x_i}(\tau)\,d\tau \qquad \text{(IV-28)}$$

the left-hand side of Equation (IV-27) may be written explicitly as:

$$\langle \underline{g}(\underline{y}), B\,\underline{\delta x_i} \rangle = \int_{t_0}^{T} \underline{g}[\underline{y}(t)] \int_{t_0}^{t} \Phi(t,\tau)H_i(\tau)\,\underline{\delta x_i}(\tau)\,d\tau\,dt \qquad \text{(IV-29)}$$

For a physical system the impulse response matrix $\Phi(t,\tau)H_i(\tau)$ is identically equal to zero for $\tau > t$; therefore the upper limit of the inner integral in the right-hand side of Equation (IV-29) may be changed to T:

$$\langle \underline{g}(\underline{y}), B\,\underline{\delta x_i} \rangle = \int_{t_0}^{T} \underline{g}[\underline{y}(t)] \int_{t_0}^{T} \Phi(t,\tau)H_i(\tau)\,\underline{\delta x_i}(\tau)\,d\tau\,dt \qquad \text{(IV-30)}$$

Next the order of integration is changed resulting in:

$$\langle \underline{g}(\underline{y}), B\,\underline{\delta x_i} \rangle = \int_{t_0}^{T}\int_{t_0}^{T} \underline{g}[\underline{y}(t)]\,\Phi(t,\tau)H_i(\tau)\,\underline{\delta x_i}(\tau)\,dt\,d\tau \qquad \text{(IV-31)}$$

Equation (IV-31) may also be written as:

$$\langle \underline{g}(\underline{y}), B\,\underline{\delta x_i} \rangle = \int_{t_0}^{T} \left\{ H_i(\tau)\,\underline{\delta x_i}(\tau) \right\}^T \int_{t_0}^{T} \left\{ \underline{g}[\underline{y}(t)]\,\Phi(t,\tau) \right\}^T dt\,d\tau \qquad \text{(IV-32)}$$

The variables of integration t and τ are interchanged:

$$\langle \underline{g}(\underline{y}), B\,\underline{\delta x}_i \rangle = \int_{t_o}^{T} \underline{\delta x}_i^{\,T}(t) H_i^{\,T}(t) \int_{t_o}^{T} \Phi^T(\tau,t)\underline{g}^T[\underline{y}(\tau)]d\tau dt \qquad \text{(IV-33)}$$

It is recognized that the adjoint differential operator $B^*\underline{g}(\underline{y})$ is given by:

$$B^*\underline{g}(\underline{y}) = H_i^T(t) \int_{t}^{T} \Phi^T(\tau,t)\underline{g}^T[\underline{y}(\tau)]d\tau \qquad \text{(IV-34)}$$

since the lower limit of integration may be changed from t_o to t ($\Phi(\tau,t) \equiv 0$ for $\tau < t$ for a physical system).

IV - 4. Physical Interpretation of the Adjoint Differential Operator.

In order to simplify the exposition a scalar version of Equation (IV-20) will be considered, i.e.

$$\frac{d}{dt} \delta y = F(t)\,\delta y + H(t)\,\delta x \qquad \text{(IV-35)}$$

Let $k(t,\tau)$ be the response of this linear system to a unit impulse $\delta(t-\tau)$ and let an operator B be defined by

$$B = \int_{o}^{t} k(t,\tau)d\tau \qquad \text{(IV-36)}$$

In the interval (0,T) the adjoint operator corresponding to Equation (IV-36) is given by:

$$B^* = \int_t^T k(\tau,t)d\tau \qquad \text{(IV-37)}$$

and

$$B^*g(y) = \int_t^T k(\tau,t)g[y(\tau)]d\tau \qquad \text{(IV-38)}$$

The solution to the differential Equation (IV-35), dropping the initial condition term, is given by:

$$\delta y(x, \delta x) = \int_0^t k(t,\tau)\, \delta x(\tau)d\tau \qquad \text{(IV-39)}$$

A reverse operator R is defined by the relation:

$$R\left\{x(t)\right\} = x(T-t), \quad t \in (0,T) \qquad \text{(IV-40)}$$

where x(t) belongs to the interval (0,T). Using definition (IV-40) in conjunction with Equation (IV-39) results in

$$R\left\{\delta y(x, \delta x)\right\} = \int_0^{T-t} k(T-t,\tau)\, \delta x(\tau)d\tau \qquad \text{(IV-41)}$$

This last equation, with a change of the variable of integration from τ to $T-\tau$, is written as:

$$R\left\{\delta y(x,\delta x)\right\} = \int_t^T k(T-t,T-\tau)\,\delta x(T-\tau)\,d\tau \qquad \text{(IV-42)}$$

But

$$\delta x(T-\tau) = R\left\{\delta x(\tau)\right\} \quad \text{with} \quad \tau \in (0,T) \qquad \text{(IV-43)}$$

Introducing a specific variation:

$$\overline{\delta x}(t) = R\left\{ g[y(t)]\right\} \qquad \text{(IV-44)}$$

and since $R\left\{R\left\{x(t)\right\}\right\} = x(t)$ (Property 1 Equation III-81) it follows that:

$$R\left\{\delta y(x,\overline{\delta x})\right\} = \int_t^T k(T-t,T-\tau)\,g[y(\tau)]\,d\tau \qquad \text{(IV-45)}$$

Since it is assumed that the plant dynamics remain time-invariant in the interval (0,T) the coefficients F and H of the differential equation (IV-35) do not vary in the same interval and the expressions on the right-hand side of Equations (IV-38) and (IV-45) are identical; thus, through linearization, a first order approximation for the adjoint differential dy* is given for $t \in (0,T)$ by running in reverse the monitored output differential dy which in turn can be obtained approximately as suggested by Pearson and Sarachik [IV-1]. This approximate identification of the adjoint differential does not require knowledge of the topology of the system differential equations. If the topology is known then dy^* may be related exactly for a limited class of systems to the plant output differentials; the procedure is the same as that introduced in Chapter III for processes described by integral operator equations.

A second method for the identification of the adjoint differential operator is based on an expansion of the operator $dy^*[x,g(t)]$ in terms of biorthogonal sequences of time functions in the interval $(0,T)$. That is:

$$dy^*[x,g] = \sum_i a_i \varphi_i(t) \qquad t \in (0,T) \tag{IV-46}$$

with

$$\langle \varphi_i(t), h_j(t) \rangle = \delta_{ij} \tag{IV-47}$$

Theoretical aspects of this kind of representation have been introduced in Chapter III Section 3. An expression for the unknown coefficients a_i (in terms of measurable differentials) is obtained by substitution of the series representation (IV-46) into the defining equation for the adjoint. Thus:

$$a_i = \frac{1}{T} \int_0^T g[x(t)]dy[x,h_i]dt \tag{IV-48}$$

The choice of the time functions $\varphi_i(t)$ is based here as in

Chapter III on the Gram-Schmidt orthogonalization process which utilizes the fewest number of terms and closely approximates the adjoint operator.

IV - 5. Implementation of Solution.

Having determined the adjoint differential operator the evaluation of the Lagrange-multiplier λ is incorporated in the total implementation scheme which is identical to the one presented in Chapter III, Section 8 for the integral operator case. The solution, that is, proceeds as follows:

1. The unconstrained problem is solved numerically, i.e. Equation (IV-17) with $\lambda = 0$ is solved iteratively for the control signal $\underline{x}^*(t)$ starting with an arbitrary initial input $\underline{x}_o(t)$. Anyone of the iteration algorithms presented in the previous chapter may be used for the evaluation of $\underline{x}^*(t)$. Then r_o is determined by:

$$\| \underline{x}^*(t) \| = r_o \tag{IV-49}$$

2. If $R \geq r_o$ the solution to the constrained problem is given by the unconstrained solution.
3. If $R < r_o$ then operation on the boundary of the constraint will dictate that:

$$\| \underline{x} \| = R \tag{IV-50}$$

and the Lagrange-multiplier is evaluated from Equation (IV-50). Steps one, two and three are repeated at each stage of the identification and optimization procedure to insure that the input constraint is always satisfied. The discussion pertaining to the implementation scheme as presented in Chapter III carries over to the present case of plants described by a set of nonlinear differential equations.

BIBLIOGRAPHY

IV-1. A.E. Pearson and P.E. Sarachik, " On the Formulation of Adaptive Optimal Control Problems ", Proc. of JACC, pp. 13 - 22, 1964.

IV-2. P.E. Sarachik and G.M. Kranc, " On Optimal Control of Systems with Multinorm Constraints ", Proc. of the 2^{nd} International Congress of IFAC, 1963.

IV-3. E. Coddington and N. Levinson, " Theory of Ordinary Differential Equations ", (book) Mc Graw-Hill Book Co., New York, 1955.

CHAPTER V

DISCRETE SYSTEMS

In the previous two chapters an approach to the problem of the adaptive optimal control of nonlinear systems was presented. The plant was described by a nonlinear continuous integral or differential operator mapping input time functions into output time functions.

The complexity of the operations will most certainly require the use of a digital computer to implement the adaptive control scheme. The computer through a digital - to - analog converter delivers piecewise constant control signals, the levels of which may be changed at discrete instants of time only. If the output is monitored at these discrete-time instants the plant can be thought of as being characterized by a discrete operator mapping discrete input functions into discrete output functions. Bigelow [V-1] formulated the sampled version of the continuous case using the methods of functional analysis. Sarachik [V-2, 3] followed a more straightforward method of analysis using calculus of variations techniques in a finite dimensional space. In this thesis a more general problem is considered.

With the tools of functional analysis, the extension is made to the multiple input-output discrete case and a solution is presented for the conditional optimization problem. A bound is placed on the input time function and the Lagrange-multiplier technique is used to formulate the modified performance functional. Finally, an ℓ_p-space analysis is found to be necessary for the inclusion of a wide class of performance criteria.

V - 1. The Discrete Single Input-Output System.

It is expedient first to develop the fundamental concepts in terms of signals with only a single component. This approach, as in Chapter III, will facilitate a better understanding of the basic assumptions, the notation and the method of formulation and solution of the problem. It is assumed that the input x and the output y are discrete time signals. Furthermore, the input will be assumed to be unconstrained with the treatment of the case with an input constraint deferred for a later section. The results here are similar to those reported by Bigelow [V-1].

The relationship between the plant output y at time t_n, namely $y(t_n)$, and the discrete-time input x will be denoted by:

$$y(t_n) = A_n[x] \qquad n = 1, 2, \ldots N \tag{V-1}$$

where A_n is generally a nonlinear summing operator having a weak differential at each instant of time t_n. The discrete levels corresponding to $n = 1$, $n = 2, \ldots n = N$ extend through the whole optimization interval $(0,T)$, starting with $t_1 = 0$ and ending with $t_N = T$, and this interval is equally subdivided. It is assumed that the plant is a combination of linear dynamics and zero-memory nonlinearities; for the identification process the number of pure accumulators in the linear part must be known a priori, and also the form of the nonlinearity, whether even or odd, must be determined. The following factors are given or assumed:

a) The performance functional, which for the case under consideration will be of the form:

$$P[x] = \sum_{n=1}^{N} G[x(t_n),\ y(t_n),\ y^d(t_n)] \qquad \text{(V-2)}$$

where, by assumption, G is a given twice differentiable function in x and y.

b) The desired output state $y^d(t_n)$ taken to be constant, and

c) The control signal set X. The problem is to determine a control signal $x^o(t_n)$, $n = 1, 2, \ldots N$, which belongs to X and minimizes $P[x]$.

The necessary condition that $P[x]$ attain an extremum with respect to x is that the first differential $dP[x,h]$ vanish; $h(t_n)$ being an arbitrary amplitude variation of x ($h(t_n)$ is also specified at discrete instants of time t_n, $n = 1, 2, \ldots N$). From Equation (V-2) this differential will be:

$$dP\lfloor x,h\rfloor = \sum_{n=1}^{N}\left\{ \frac{\partial G}{\partial x(t_n)} h(t_n) + \frac{\partial G}{\partial y(t_n)} dA\lfloor x(t_n),h(t_n)\rfloor\right\} \tag{V-3}$$

It is more convenient to formulate the necessary condition of optimality in a form independent of the arbitrary variation h. With this as a motivation the adjoint differential operator $dA^*\lfloor x(t_n),g(t_n)\rceil$ is defined by the inner product relationship:

$$\sum_{n=1}^{N} g(t_n)\ dA\lfloor x(t_n),h(t_n)\rfloor = \sum_{n=1}^{N} h(t_n)\ dA^*\lfloor x(t_n),g(t_n)\rceil \tag{V-4}$$

Recognizing that:

$$h(t_n) = \frac{\partial G}{\partial y(t_n)}$$

and substituting Equation (V-4) into Equation (V-3) a relative extremum of $P\lfloor x\rceil$ exists if

$$\sum_{n=1}^{N} h(t_n)\left\{ \frac{\partial G}{\partial x(t_n)} + dA^*\lfloor x(t_n), \frac{\partial G}{\partial y(t_n)} \rceil\right\} = 0 \tag{V-5}$$

Since by assumption $h(t_n)$ is arbitrary in the interval $n = 1$ to $n = N$, it follows from Equation (V-5) that $P[x]$ attains a relative extremum when

$$\frac{\partial G}{\partial x(t_n)} + dA^*\left[x(t_n), \frac{\partial G}{\partial y(t_n)}\right] = 0 \qquad \text{for } n = 1, 2, \ldots N \tag{V-6}$$

An explicit solution for the control signal $x(t_n)$ will require:

a) The identification of the adjoint differential operator $dA^*[x(t_n), g(t_n)]$ at all instants of time t_1, t_2, ... t_n;

b) some iterative scheme for the solution of the nonlinear operator Equation (V-6).

Two different methods for the identification of dA^* are given in Chapter III. They will be utilized here for the identification of the dicrete adjoint operator. Both methods presume knowledge of the differential $dA[x(t_n), h(t_n)]$, and the latter is conveniently approximated by:

$$dA[x(t_n), h(t_n)] = \frac{A[x(t_n) + \gamma h(t_n)] - A[x(t_n)]}{\gamma} \tag{V-7}$$

where γ is a small number.

The first method, based on the properties of the reverse operator, is limited to some classes of systems and presumes knowledge of the general plant structure; it will be illustrated assuming the following form for the plant operator:

$$A[x] = y_o(t_i) + \sum_{n=1}^{i} k(t_i,t_n) \lfloor 1 - e^{-ax(t_n)} \rfloor \qquad \text{(V-8)}$$

with $k(t_i,t_n)$ an unknown kernel. The differential of $A[x]$ is given by:

$$dA[x,h] = \sum_{n=1}^{i} k(t_i,t_n) a e^{-ax(t_n)} h(t_n) \qquad \text{(V-9)}$$

and the adjoint differential operator $dA^*[x(t_i)\ g(t_i)]$ is defined by the inner product relationship expressed by Equation (V-4) which is repeated here for convenience:

$$\sum_{i=1}^{N} g(t_i)\ dA[x(t_i,h(t_i)] = \sum_{i=1}^{N} h(t_i)\ dA^*[x(t_i),g(t_i)] \qquad \text{(V-10)}$$

Substituting Equation (V-9) into Equation (V-10) we obtain:

$$\sum_{i=1}^{N} g(t_i) \sum_{n=1}^{i} k(t_i,t_n) a e^{-ax(t_n)} h(t_n) = \sum_{i=1}^{N} h(t_i) dA^*[x(t_i),g(t_i)] \qquad \text{(V-11)}$$

Since $k(t_i,t_n) = 0$ for $n > i$ for a physical system we may write Equation (V-11) in the form:

$$\sum_{i=1}^{N} g(t_i) \sum_{n=1}^{N} k(t_i,t_n)\ ah(t_n)e^{-ax(t_n)} = \sum_{i=1}^{N} h(t_i)dA^*\lfloor x(t_i),g(t_i)\rfloor \tag{V-12}$$

According to the properties of the bilinear inner product relationship we may interchange the order of summation resulting in:

$$\sum_{n=1}^{N} ah(t_n)e^{-ax(t_n)} \sum_{i=1}^{N} k(t_i,t_n)g(t_i) = \sum_{i=1}^{N} h(t_i)dA^*\lfloor x(t_i),g(t_i)\rfloor \tag{V-13}$$

An interchange of the running indeces n and i results in:

$$\sum_{i=1}^{N} ah(t_i)e^{-ax(t_i)} \sum_{n=1}^{N} k(t_n,t_i)g(t_n) = \sum_{i=1}^{N} h(t_i)dA^*\lfloor x(t_i),g(t_i)\rfloor \tag{V-14}$$

But $k(t_n,t_i) = 0$ for $t_i > t_n$ (a physical system is again under consideration), so:

$$\sum_{i=1}^{N} h(t_i)ae^{-ax(t_i)} \sum_{n=i}^{N} k(t_n,t_i)g(t_n) = \sum_{i=1}^{N} h(t_i)dA^*\lfloor x(t_i),g(t_i)\rfloor \tag{V-15}$$

From the above equation dA* is recognized to be:

$$dA^*[x(t_i)\ g(t_i)] = ae^{-ax(t_i)} \sum_{n=i}^{N} k(t_n, t_i) g(t_n) \qquad (V-16)$$

In general with $N[x]$ representing a twice differentiable zero-memory nonlinear device, and

$$A[x] = y_o(t_i) + \sum_{n=1}^{i} k(t_i, t_n)\ N[x(t_n)] \qquad (V-17)$$

the adjoint differential operator takes the form:

$$dA^*[x(t_i), g(t_i)] = \frac{dN}{dx(t_i)} \sum_{n=i}^{N} k(t_n, t_i) g(t_n) \qquad (V-18)$$

Introducing a discrete inverse operator R by the relationship:

$$R\left\{x(t_i)\right\} = x(t_{N-i}) \qquad (V-19)$$

where t_i is an element of the discrete time interval t_1 to t_n and $x(t_i)$ belongs to this interval, and applying this reverse operator to the expression for the differential of $A[x]$, with a change of the summation index and proper attention given to discrete differences from the initial point, results in:

$$R\left\{dA[x,h]\right\} = \sum_{n=i}^{N} k(t_{N-i}, t_{N-n}) ae^{-ax(t_{N-n})}\ h(t_{N-n}) \qquad (V-20)$$

But

$$ae^{-ax(t_{N-n})} \quad h(t_{N-n}) = R\left\{ ae^{-ax(t_n)} \; h(t_n)\right\} \tag{V-21}$$

With a partisular variation $\bar{h}(t_n)$, given by:

$$\bar{h}(t_n) = \frac{R\left\{g(t_n)\right\}}{ae^{-ax(t_n)}} \tag{V-22}$$

and utilizing property 1, Equation (III-81), of the reverse operator, Equation (V-20) can be written:

$$R\left\{ dA[x,\bar{h}]\right\} = \sum_{n=1}^{N} k(t_{N-1},t_{N-n})g(t_n) \tag{V-23}$$

If the system under consideration has a kernel which is not varying in the time interval (0,T) then Equation (V-23) takes the form:

$$R\left\{ dA[x,\bar{h}]\right\} = \sum_{n=1}^{N} k(t_n,t_1)g(t_n) \tag{V-24}$$

It is recognized that the discrete adjoint differential operator $dA^*[x(t_1),g(t_1)]$ is now given by:

$$dA^*[x(t_1),g(t_1)] = ae^{-ax(t_1)} \; R\left\{ dA[x,\bar{h}]\right\} \tag{V-25}$$

For the more general plant structure of Equation (V-17)

$dA^*[x(t_i),g(t_i)]$ is written explicitly as:

$$dA^*[x(t_i),g(t_i)] = \frac{dN}{dx(t_i)} \; R\left\{ dA[x,\bar{h}] \right\} \tag{V-26}$$

Experimentally the discrete levels of the adjoint operator are found by reversing the order of the discrete levels of $dA[x,\bar{h}]$ and multiplying each by the corresponding value of $\frac{dN}{dx(t_i)}$.

The second method for the identification of the adjoint operator is based on Kulikowski's expansion technique; he suggested an expansion of $dA^*[x,g]$ in terms of orthonormal time functions for plants described by continuous operator equations. Presently, we may express $dA^*[x,g]$ as:

$$dA[x,g] = \sum_{\ell=1}^{\infty} a_\ell \, h_\ell \,(t_n) \tag{V-27}$$

with

$$\frac{1}{N} \sum_{n=1}^{N} h_i(t_n)h_j(t_n) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{V-28}$$

Substitution of Equation (V-27) into the defining equation for the adjoint operator, Equation (V-4), results in:

$$\sum_{n=1}^{N} g(t_n)dA[x(t_n),h(t_n)] = \sum_{n=1}^{N} h(t_n) \sum_{\ell} a_\ell \, h(t_n)h_\ell \,(t_n) \tag{V-29}$$

or:

$$\sum_{n=1}^{N} g(t_n)dA[x(t_n),h(t_n)] = \sum_{n}\sum_{\ell} a_{\ell}\, h(t_n)h_{\ell}(t_n) \qquad \text{(V-30)}$$

From the orthogonality condition and Equation (V-30) above follows immediately that:

$$a_{\ell} = \frac{1}{N}\sum_{n=1}^{N} g(t_n)dA[x(t_n),h_{\ell}(t_n)] \qquad \text{(V-31)}$$

The discrete time functions $h_{\ell}(t_n)$ must form an orthonormal set and can be chosen conveniently as the unit vectors along the orthogonal axes, i.e.

$$h_{\ell}(t_n) = (0,\ 0,\ \ldots,\ 1,\ 0,\ \ldots)$$

where the value of 1 is assigned to the ℓ^{th} element.

Finally, the discrete nonlinear operator equation (V-6) may be solved explicitly for the control variable $x(t_n)$ utilizing a discrete version of any of the iteration algorithms discussed in Chapter III.

The conditional optimization problem is considered next. The plant is again supposed to consist of a single input and a single output and the analysis is carried out in an ℓ_p-space. As it was pointed out by Sarachik [V-3], the input and output functions are actually elements of finite dimensional spaces; this realization leads to simplified results regarding the identification problem and the iterative solution of the nonlinear operator equation [V-6]. Yet, it is found to be convenient to formulate the solution with both the input and output functions assumed to be elements in ℓ_p-space.

V - 2. Conditional Optimization of Discrete Systems.

For the single input-output case the conditional optimization problem is formulated in the following way: The functional $\|y^d - A[x]\|_{p_1}$ is to be minimized subject to the constraint

$$\|x\|_{p_2} \leq R \tag{V-32}$$

where R is a given positive constant and $p_1, p_2 \geq 1$. A regulator problem is presently under consideration. A wide class of performance criteria may be formulated using the same approach and depending on the plant requirements; the analysis proceeds along similar lines. The norm of a discrete time function $x(t_n)$ in ℓ_p-space is defined by:

$$\|x(t_n)\|_p \triangleq \left(\sum_{n=1}^{N} |x(t_n)|^p \right)^{1/p} \tag{V-33}$$

The augmented performance functional $P[x]$ is formed:

$$P[x] = \lambda \|x\|_{p_2} + \|y^d - A[x]\|_{p_1} \qquad \text{(V-34)}$$

where λ is the Lagrange-multiplier and will be determined from the input constraint condition. Different types of physical constraints could be imposed by choosing different values for the parameter p_2. If $p = 2$ in Equation (V-33) the "energy" or "power" of the input signal is constrained; $p = 1$ corresponds to an area constraint and for $p = \infty$ we obtain the set of signals bounded in magnitude. It is our purpose to determine a sequence of input signals which will minimize the performance functional. The necessary condition for the existence of an extremum of $P[x]$ is again expressed by the vanishing of the differential of the performance functional. The first differential of $P[x]$ will be evaluated from the defining equation:

$$dP[x,h] = \left. \frac{dP[x(t_n) + \gamma h(t_n)]}{d\gamma} \right|_{\gamma = 0} \qquad \text{(V-35)}$$

Substitution of $x(t_n) + \gamma h(t_n)$ for $x(t_n)$ in Equation (V-34) and differentiation with respect to γ yields:

$$dP[x,h] = \lambda \frac{1}{p_2} \left(\sum_{n=1}^{N} |x(t_n)|^{p_2} \right)^{\frac{1}{p_2}-1} p_2 \sum_{n=1}^{N} |x(t_n)|^{p_2-1} \mathrm{sgn}[x(t_n)]\, h(t_n)$$

$$- \frac{1}{p_1} \left(\sum_{n=1}^{N} |y^d - A[x]| \right)^{\frac{1}{p_1}-1} p_1 \sum_{n=1}^{N} |y^d - A[x]|^{p_1-1} \mathrm{sgn}(y^d - A[x(t_n)] \quad)] \tag{V-36}$$

with $\mathrm{sgn}[z(t_n)]$ defined by:

$$\mathrm{sgn}[z(t_n)] = \begin{cases} +1 & \text{for} \quad z(t_n) > 0 \\ -1 & \text{for} \quad z(t_n) < 0 \end{cases} \tag{V-37}$$

As in the previous section the adjoint differential operator will be introduced, this being motivated by the desire of expressing the necessary condition of optimality in a form which does not contain the arbitrary variation $h(t_n)$. For a more concise formulation let:

$$|y^d - A[x(t_n)]|^{p_1-1} \mathrm{sgn}(y^d - A[x(t_n)]) = g[x(t_n)] \tag{V-38}$$

Then $dA^*[x(t_n), g(t_n)]$, the adjoint differential operator, is defined by Equation (V-10) which is repeated here for convenience:

$$\sum_{n=1}^{N} g(t_n)\, dA[x(t_n), h(t_n)] = \sum_{n=1}^{N} h(t_n)\, dA^*[x(t_n), g(t_n)] \tag{V-39}$$

Substituting Equation (V-39) into Equation (V-36) the necessary condition for the existence of an extremum can be rewritten as:

$$dP[x,h] = \sum_{n=1}^{N} h(t_n) \left\{ \frac{\lambda}{\|x(t_n)\|_{p_2}^{p_2-1}} |x(t_n)|^{p_2-1} \operatorname{sgn}[x(t_n)] \right.$$

$$\left. - \frac{1}{\|y^d - A[x(t_n)]\|_{p_1}^{p_1}} \, dA^*[x(t_n), g(t_n)] \right\} \tag{V-40}$$

or:

$$|x(t_n)|^{p_2-1} \operatorname{sgn}[x(t_n)] = \frac{1}{\lambda} \frac{\|x(t_n)\|_{p_2}^{p_2-1}}{\|y^d A[x(t_n)]\|_{p_1}^{p_1-1}} \tag{V-41}$$

since $h(t_n)$ is arbitrary in the interval $n = 1$ to $n = N$. Equation (V-41) is similar to the expression derived for the continuous case in Chapter III. For different values of p_1 and p_2 Equation (V-41) may be solved explicitly for the control signal $x(t_n)$ utilizing a discrete version of anyone of the iteration schemes described previously.

Finally, intuitive considerations in most physically motivated problems dictate that the norm of the input $x(t_n)$ lies on the boundary of the constraint, or:

$$\|x\|_{p_2} = R \qquad \text{(V-42)}$$

From Equation (V-42) the Lagrange-multiplier λ is evaluated. The implementation scheme involves exactly the same steps as those described in Chapters III and IV for continuous systems.

V - 3. Multiple Input - Output Discrete Systems.

Consider the system shown in Figure (V-1). The nonlinearities $N_1, \ldots N_m$ and N are zero-memory and the linear parts $L_1, \ldots L_m$ are of type zero, i.e. the plant possesses no pure accumulators. The particular plant structure considered here facilitates a simple presentation of the basic ideas; more complicated plant structures may be studied with suitable extensions of the identification and optimization techniques.

There are m inputs to the plant and k outputs. The input vector $\underline{x}(t_n)$, at the instant t_n, is defined to be the column vector:

$$\underline{x}(t_n) = \text{col}\left\{ x_1(t_n), x_2(t_n) \ldots x_m(t_n) \right\} \qquad \text{(V-43)}$$

Similarly the output vector $\underline{y}(t_n)$ is the column vector:

$$\underline{y}(t_n) = \text{col}\left\{ y_1(t_n),\ y_2(t_n), \ldots\ y_k(t_n) \right\} \tag{V-44}$$

The i^{th} component of the output vector is related to the input vector through the nonlinear plant operator A according to the relationship:

$$y_i(t_n) = A_i[x_1(t_n),\ x_2(t_n), \ldots\ x_m(t_n)] \qquad i = 1,\ 2,\ \ldots\ k \tag{V-45}$$

Or in a more general vector form:

$$\underline{y}(t_n) = \underline{A}[\underline{x}(t_n)] \qquad \text{for} \quad n = 1,\ 2,\ \ldots\ N \tag{V-46}$$

The weak differential of the discrete vector operator $\underline{A}[\underline{x}]$ is by definition:

$$d\underline{A}[\underline{x},\underline{h}] = \frac{\partial}{\partial \gamma}\underline{A}[\underline{x}(t_n) + \gamma\,\underline{h}(t_n)] \tag{V-47}$$

where $\underline{h}(t_n)$ is a column vector of m arbitrary components belonging to the input function set X. The differential $d\underline{A}[\underline{x},\underline{h}]$ may be considered as a k x m matrix of derivative operators operating on the vector $\underline{h}$, or:

$$dA[\underline{x},\underline{h}] = \begin{bmatrix} \begin{bmatrix} \dot{A}_{11} & \dot{A}_{12} \cdot \cdot \cdot & \dot{A}_{1m} \\ \vdots & & \vdots \\ \dot{A}_{k1} & \dot{A}_{k2} \cdot \cdot \cdot & \dot{A}_{km} \end{bmatrix} & \begin{matrix} h_1 \\ \vdots \\ h_m \end{matrix} \end{bmatrix} \tag{V-48}$$

where $\dot{A}_{ij} = \dot{A}_{ij}[\underline{x}]$ represents the derivative operator of $A_i[\underline{x}]$ corresponding to a variation in x_j only, i.e.

$$A_{ij}[\underline{x}]h_j = \lim_{\gamma \to 0} \frac{A_i[x_1, x_2 \ldots x_j + \gamma h_j, \ldots x_m] - A_i[\underline{x}]}{\gamma} \tag{V-49}$$

with $i = 1, 2, \ldots k$ and $j = 1, 2, \ldots m$

The i^{th} differential of the vector operator $\underline{A}[\underline{x}]$ follows from Equation (V-48) as:

$$dA_i[\underline{x},\underline{h}] = \sum_{j=1}^{m} \dot{A}_{ij}[\underline{x}]g_j \tag{V-50}$$

$i = 1, 2, \ldots k$

With the same problem in mind as treated in Section 1 of this chapter for the single input-output case, the performance functional for the multi-dimensional system will be:

$$P[\underline{x}] = \sum_{n=1}^{N} G[\underline{x}(t_n),\ \underline{y}(t_n)\ \underline{y}^d(t_n)] \qquad \text{(V-51)}$$

And the necessary condition for the existence of an extremum is expressed as:

$$dP[\underline{x},\underline{h}] = \sum_{n=1}^{N} \left\{ \frac{\partial G}{\partial \underline{x}(t_n)} \quad \underline{h}(t_n) + \frac{\partial G}{\partial \underline{y}(t_n)} \, dA[\underline{x}(t_n),\underline{h}(t_n)] \right\} = 0 \qquad \text{(V-52)}$$

where $\frac{\partial G}{\partial \underline{x}}$ is an m-dimensional vector and $\frac{\partial G}{\partial \underline{y}}$ is a k-dimensional vector with components respectively:

$$\left\{ \frac{\partial G}{\partial \underline{x}(t_n)} \right\}_i = \frac{\partial}{\partial x_i(t_n)} \quad G[\underline{x}(t_n),\underline{y}(t_n),\underline{y}^d(t_n)]$$

$$\left\{ \frac{\partial G}{\partial \underline{y}(t_n)} \right\}_j = \frac{\partial}{\partial y_j(t_n)} \quad G[\underline{x}(t_n),\underline{y}(t_n),\underline{y}^d(t_n)]$$

$i = 1,\ 2,\ \ldots\ m \quad j = 1,\ 2,\ \ldots\ k$

and $n = 1,\ 2,\ \ldots\ N$

For the development of the adjoint differential operator, define an m x k matrix of derivative operators $A'^*[\underline{x}(t_n)]$ by the inner product relationship:

$$\sum_{n=1}^{N} \underline{g}(t_n) \left\{ A'\underline{x}(t_n)]\ \underline{h}(t_n) \right\} = \sum_{n=1}^{N} \underline{h}(t_n) \left\{ A'^*[\underline{x}(t_n)]\underline{g}(t_n) \right\} \qquad \text{(V-53)}$$

Substituting Equation (V-53) into Equation (V-52) we obtain:

$$dA[\underline{x},\underline{h}] = \sum_{n=1}^{N} \underline{h}(t_n) \left\{ \frac{\partial G}{\partial \underline{x}(t_n)} + A'^*[\underline{x}(t_n)] \frac{\partial G}{\partial \underline{y}(t_n)} \right\} = 0$$

(V-54)

Or, since $\underline{h}$ is an arbitrary vector function

$$\frac{\partial G}{\partial \underline{x}(t_n)} + A'^*[\underline{x}(t_n)] \frac{\partial G}{\partial \underline{y}(t_n)} = 0 \qquad \text{(V-55)}$$

$$n = 1, 2, \ldots N$$

The elements of matrix $A'^*[\underline{x}(t_n)]$ are recognized to be the adjoints of the corresponding elements of the matrix $A'[\underline{x}(t_n)]$ transposed. In matrix form, designating the adjoint of A'_{ij} by A'^*_{ij}:

$$A'^*[\underline{x}(t_n)]\underline{g}(t_n) = \begin{bmatrix} \begin{bmatrix} A'^*_{11} & A'^*_{21} & \cdots & A'^*_{k1} \\ \vdots & & & \vdots \\ A'^*_{1m} & A'^*_{2m} & \cdots & A'^*_{km} \end{bmatrix} & \begin{matrix} g_1 \\ \vdots \\ g_k \end{matrix} \end{bmatrix} \qquad \text{(V-56)}$$

The techniques for the identification of the adjoint are the same as discussed for the one-dimensional case with each component of the input vector perturbed separately for the estimation of the plant differentials. If the input and output spaces are considered to be finite-dimensional then the operator $A'^{*}[\underline{x}(t_n)]$ is given by the transpose of the matrix for the operator $A'[\underline{x}(t_n)]$ (See Reference (V-4), p. 103).

The discrete operator Equation (V-55) may be solved explicitly for the control vector $\underline{x}(t_n)$. The principle of contraction mapping, Newton's method or the technique by Altman could be applied for the solution. For example, if Equation (V-55) is written as:

$$\underline{Q}[\underline{x}(t_n)] = 0 \qquad \text{(V-57)}$$

Altman's method generates the sequence of input signals through the iteration:

$$\underline{x}_{m+1}(t_n) = \underline{x}_m(t_n) - \frac{\| \underline{Q}[\underline{x}_m(t_n)] \|_p^p}{p \| \underline{f}[\underline{x}_m(t_n)] \|_q} \quad \underline{y}(t_n) \qquad \text{(V-58)}$$

$$n = 1, 2, \ldots N \quad \text{and} \quad \frac{1}{p} + \frac{1}{q} = 1$$

with the components of $\underline{f}$ and $\underline{y}$ defined as in Section (III-7) for the continuous single-dimensional case.

The multi-dimensional conditional optimization problem follows the formulation presented in Section (V-2) with suitable vector notation, and for its solution similar steps as those introduced above for the multivariable unconstrained problem.

Computer flow diagrams for the digital implementation of the adaptive control algorithms are presented in the next chapter where specific applications are discussed.

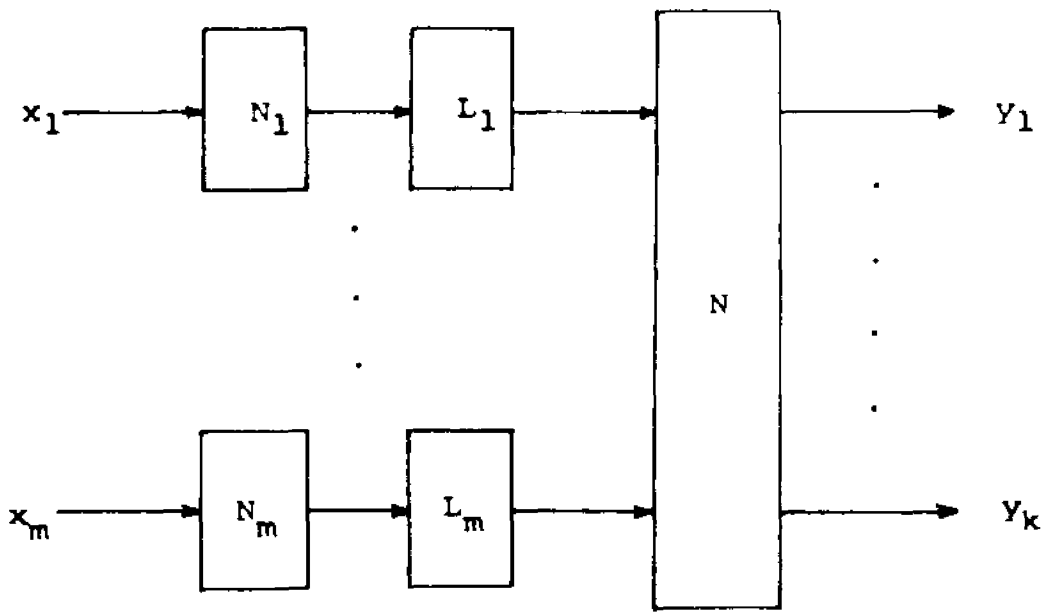


FIGURE V-1

BIBLIOGRAPHY

V-1. S.C. Bigelow, " Adaptive Control of Nonlinear Sampled Data Systems ", D. Engr. Sci. Dissertation, Columbia University, 1963.

V-2. P.E. Sarachik, " Identification of the Steady-State Operator for Discrete Self-Optimizing Systems ", Trans. IEEE, AC - 10, No. 1, Vol. 83, 1965.

V-3. P.E. Sarachik, " An Approach to the Design of Non-linear Discrete Self Optimizing Systems ", Third Symposium on Discrete Adaptive Processes, 1964.

V-4. A.N. Kolmogorov and S.V. Fomin, " Elements of the Theory of Functions and Functional Analysis ", Vol. 1, Metric and Normed Spaces, Graylock Press, Rochester, New York, 1957.

CHAPTER VI

CASE STUDIES

In this chapter three case studies are presented. The examples were chosed for the simplicity of the plant equations, the ease of mathematical formulation and in some instances the availability of an explicit solution; they illustrate basic concepts avoiding unnecessary details. It is assumed that an exact description of the plant equations is available. This is necessary for the digital implementation of the adaptive scheme since a continuous record of the system output is needed. If the scheme were to be applied to an actual process, then direct monitoring of the process output would suffice. The approach followed here offers a number of distinct advantages: It is possible in some cases to evaluate directly the plant differentials thus furnishing a means for comparison with the approximate technique and providing useful information regarding the choice of values for the parameter γ. Similarly the identification of the adjoint differential may be accomplished in a dual fashion affording the same kind of comparison. In one case a mathematical transformation of the system equation from the integral operator form to the differential operator description is feasible. The two approaches described in Chapters III and IV are found to lead to identical results.

At the same time with an a priori knowledge of the plant equations and a digital simulation of the overall adaptive scheme some of the procedural steps are eliminated. The initial learning period is not present here; also the initial conditions can be set to the desired values at the beginning of each period simplifying the identification problem for plants possessing one or more pure accumulators.

The main stumbling block for the mechanization of any complex or multivariable process is the computational problem associated with the search procedures presently available. First variational prodedures are found not to converge and second variational techniques are difficult to implement. A modified Newton's method was used in most cases and reasonably satisfactory results were obtained. Numerical results were obtained using an I B M 360 Model 50 computer at the Computation Center of the City College of the City University of New York. The programs were written in the FORTRAN IV language and the actual computer print-outs are included in Appendix II. In the following sections the problem formulation, digital implementation and graphical results for the three case studies are described.

VI - 1. Case Study Number 1.

As a first example the plant structure shown in Figure VI-1 is considered. The input nonlinearity N is given by:

$$N(x) = c[1-e^{-ax}] \tag{VI-1}$$

with c and a given constants specifying the nonlinear characteristics of N. In an actual process these parameters need not be known a priori, but may be determined through steady-state measurements as it is shown in Chapters I and III. Figure VI-2 shows a plot of N(x) vs. the input variable x. N(x) is zero - memory.

The linear part L of the plant is assumed to be type zero. In its simplest form it has a transfer function specified by:

$$\frac{Y}{V}(s) = \frac{\text{const.}}{s+b} \tag{VI-2}$$

where b is a constant.

The overall output time response is given by:

$$y(t) = A[x] = A_o + \int_o^t c' e^{-b(t-\tau)} [1-e^{-ax(\tau)}] d\tau \tag{VI-3}$$

For the computer solution the following values were used for the constants: $c' = 1$, $a = 0.5$, $b = 0.5$. The plant settling time is given approximately by $T_s = 5 \cdot 1/b = 10$. The optimization interval $(0,T)$ was chosen to be equal to ten time units, i.e. $T = T_s$ with the interval $(0,T)$ equally subdivided into ten subintervals.

We seek to minimize the functional:

$$J[x] = \left\| y^d - A[x] \right\|_2^2 \tag{VI-4}$$

under the constraint:

$$\left\| x \right\|_2^2 \leq R \tag{VI-5}$$

The constant level of the desired output state was set to unity and two different values for the positive constant R were used in the simulation; the first one resulted in an unconstrained optimal solution for $x(t)$ and the second forced the system to operate on the constraint boundary.

Minimization of the augmented performance criterion $P[x]$, where:

$$P[x] = \lambda \left\| x \right\|_2^2 + \left\| y^d - A[x] \right\|_2^2 \tag{VI-6}$$

and λ being the Langrange-multiplier, results in the following two implicit solutions for x(t):

For $\lambda = 0$ (unconstrained solution)

$$- dA^*[x,g] = 0 \tag{VI-7}$$

And for $\lambda > 0$ (constrained solution)

$$x(t) - dA^*[x,g] = 0 \tag{VI-8}$$

The adjoint differential $dA^*[x,g]$ is evaluated as follows: An estimate of the plant differential $dA[x,h]$ is obtained first through the relationship:

$$dA[x,h] \approx \frac{A[x+\gamma h]-A[x]}{\gamma} \tag{VI-9}$$

Values for γ utilized in the computer program range from $\gamma = 0.001$ to $\gamma = 0.1$ In an actual process operation of the adaptive scheme noise considerations will limit rhe lower bound of γ .

According to the results presented in Chapter III, the adjoint differential $dA^*[x,g]$ is given by:

$$dA^*[x,g] = e^{-ax} R \left\{ dA[x, e^{ax} R \{ g \}] \right\} \tag{VI-10}$$

with a particular input variation $\bar{h}(t)$ given by:

$$\bar{h}(t) = \frac{R\{g(t)\}}{dN/dx} \qquad \text{(VI-11)}$$

and

$$g(t) \triangleq y^d - A[x] \qquad \text{(VI-12)}$$

Thus the adjoint differential is evaluated experimentally by tabulating in reverse order the estimated values of $dA[x,\bar{h}]$ and multiplying each by the corresponding value of e^{ax}.

It is possible to determine directly the exact expressions for both the plant differential and its adjoint. That is from Equation (VI-3):

$$dA[x,h] = \int_0^t c' e^{-b(t-\tau)} a e^{-ax(\tau)} h(\tau) d\tau \qquad \text{(VI-13)}$$

and

$$dA^*[x,g] = e^{-ax(t)} \int_t^T c' e^{-b(\tau-t)} g(\tau) d\tau \qquad \text{(VI-14)}$$

The above two equations were used in a particular simulation and the results compare very well with the approximate ones derived from Equations (VI-9) (with $\gamma = 0.01$) and (VI-10) respectively.

The computer program incorporates a trapezoidal rule for the evaluation of the integrals. It is also possible to use other available and more accurate integration subroutines whenever this is deemed necessary.

All three iteration techniques described in Chapter III were utilized in the search for the optimal value of the control signal x(t).

It was found that the expressions (VI-7) and (VI-8) are <u>not</u> contractions with a Lipschitz constant of 0.5.

Altman's method was applied in the following way: The iterations of Equation (III-122) are to be implemented. Starting with the expression for the augmented performance criterion, Equation (VI-6), the gradient of $P\lfloor x\rfloor$ is evaluated to be:

$$\nabla P\lfloor x\rfloor = 2\,x - 2dA^*\lfloor x,g\rfloor \qquad \text{(VI-15)}$$

with

$$g = y^d - A\lfloor x\rfloor \qquad \text{(VI-16)}$$

Next, the second differential of $P\lfloor x\rfloor$ is calculated:

$$d^2P[x,h,h] = 2\int_0^T \left\{ \lambda h^2 + (dA[x,h])^2 - (y^d - A[x])d^2A[x,h,h] \right\} dt$$

(VI-17)

But

$$d^2A[x,h,h] = -a\,dA[x,h^2] \qquad \text{(VI-18)}$$

and utilizing the defining equation for the adjoint differential, the second differential of the augmented performance criterion is written as:

$$d^2P[x,h,h] = 2\int_0^T \left\{ \lambda h^2 + (dA[x,h])^2 \right\} dt + 2a\int_0^T h^2 dA^*[x, y^d - A[x]]\, dt \qquad \text{(VI-19)}$$

With a fixed input variation equal to the gradient of $P[x]$, as dictated by the iteration equation, the above expression for $d^2P[x,h,h]$ takes the form:

$$d^2P[x,\nabla P[x],\nabla P[x]] = 2\int_0^T \left\{ \lambda \nabla P[x])^2 + (dA[x,\nabla P[x]])^2 \right\} dt + 2a\int_0^T (\nabla P[x])^2 dA^*[x, y^d - A[x]]\, dt \qquad \text{(VI-20)}$$

Thus, the (n+1)st element of the control sequence is generated from the n^{th} element according to the relation:

$$x_{n+1} = x_n - \frac{(\nabla P[x_n], \nabla P[x_n])}{2d^2P[x_n, \nabla P[x_n], \nabla P[x_n]]} \cdot \nabla P[x_n] \qquad \text{(VI-21)}$$

The overall computer flow diagram is shown at the end of this chapter and a listing of the computer program is included in Appendix II. The results of this search procedure are shown in Figure (VI-3) (a) and (b). Figure (VI) (a) is a plot of the plant output and Figure (VI-I) (b) shows the control signal as a function of time for the unconstrained solution. The output eventually reaches the desired level.

Newton's modified method with a fixed stepsize was also used in conjunction with this case study. The results are shown in Figures (VI-4) and (VI-5). Figure (VI-4) (a) and (b) shows a plot of the system output and the control signal as functions of time for the unconstrained solution. Again the output eventually reaches the desired level. The speed of convergence depends of course on the choice of the arbitrary initial condition level for the input function and the magnitude of the fixed step-size. The program was allowed to run for 100 steps and the total execution time was 3.70 minutes. The objective output level is reached within a reasonable approximation in about 30 iteration steps. The last sampling point of each

T-time interval remains always equal to the value set at the start of the program because it can not be calculated exactly when the reverse operator method is used for the identification of the adjoint. This may be remedied by extrapolating or increasing the sampling points in the interval (0,T). The same unconstrained problem was programmed with a direct evaluation of the plant differential from Equation (VI-13) and the same iteration step-size. The total execution time was 2.17 minutes and the graphical results are identical with the ones presented in Figure (VI-4).

Figure (VI-5) (a) and (b) shows the results of a computer run with a severe enough input constraint to cause the control signal to operate on the boundary of the constraint. The iteration step-size has been increased from the previous case by an order of magnitude without affecting the convergence properties of the scheme. The total execution time was 2.26 minutes and the program was allowed to run for 100 steps. The constraint condition is treated as a penalty function: A search is performed for that value of the Lagrange-multiplier that forces the input function to lie on the constraint boundary. Figure (VI-6) depicts the objective function $\|y^d - A[x]\|$ vs. the step number for both Newton's modified procedure and the technique introduced by Altman. It is seen that Newton's method with a fixed step-size converges in fewer steps. The choice of the step-size value is, of course, extremely critical as far as the convergence properties of the scheme are concerned.

VI - 2. Case Study Number 2.

The plant structure considered is shown in Figure (VI-7). The general input-output relationship is given by:

$$A[x] = A_o + \int_0^t k(t,\tau)x(\tau)d\tau - \varepsilon[x(t)]^2 \qquad \text{(VI-22)}$$

where ε is a cost parameter. Both ε and the kernel $k(t,\tau)$ may be thought of as being the unknown but slowly-varing plant characteristics for the adaptive problem.

It is desired that the total plant output during a period of steady-state operation be maximized. Correspondingly, we seek to minimize the functional:

$$J[x] = - \int_0^T A[x]dt \qquad \text{(VI-23)}$$

The input norm is constrained to be less or equal to a positive constant R, or:

$$\|x\|_p \leq R \qquad \text{(VI-24)}$$

Two values for the parameter p were chosen for this example, $p = 2$ and $p \to \infty$. For $p = 2$ the gradient of the augmented performance functional is given by:

$$\nabla P[x] = 2\lambda x - dA^*[x, 1] \qquad \text{(VI-25)}$$

whereas for $p \to \infty$ the vanishing of the gradient of $P[x]$ leads to the following solution for the control signal:

$$\begin{aligned} x(t) &= \|x\|_p \to \max |x(t)| \\ \operatorname{sgn} x(t) &= \operatorname{sgn} dA^*[x, 1] \end{aligned} \qquad \text{(VI-26)}$$

Therefore

$$x(t) = R \operatorname{sgn} dA^*[x, 1] \qquad \text{(VI-27)}$$

for operation on the boundary of the constraint.

For computer simulation purposes let us assume that the plant is explicitly described by:

$$A[x] = A_o + \int_o^t a e^{-a(t-\tau)} x(\tau) d\tau - \varepsilon [x(t)]^2 \qquad \text{(VI-28)}$$

with $A_o = 0$, $a = 3/T$, $\varepsilon = 1/4$ and $T = 3$.

The first plant differential is given by:

$$dA[x,h] = \int_0^t ae^{-a(t-\tau)} h(\tau)d\tau - 2\varepsilon\, x(t)h(t) \qquad \text{(VI-29)}$$

Making use of the defining inner product form for the adjoint plant differential we obtain:

$$dA^*[x, 1] = \int_t^T ae^{-a(\tau-t)} d\tau - 2\varepsilon\, x(t) \qquad \text{(VI-30)}$$

Equation (VI-30) affords a one-step solution for the control signal x(t) in the interval (0,T). This solution corresponds to the vanishing of the gradient of $P[x]$ as it is given by Equation (VI-25) with $\lambda = 0$. The resultant optimum control signal $x^o(t)$ is shown in Figure (VI-7).

Applying the reverse operator to the plant differential, Equation (VI-29), results in:

$$R\left\{ dA[x,h]\right\} = \int_t^T ae^{-a(\tau-t)} R\left\{ h(\tau)\right\} d\tau - 2\varepsilon R\left\{ x(t)h(t)\right\} \qquad \text{(VI-31)}$$

Let $R\left\{ h(t)\right\} = 1$, or h(t) = 1 in the interval (0,T). Then

$$dA^*[x,1] = R\left\{ dA[x,1]\right\} + 2\varepsilon\left[R\left\{ x(t)\right\} - 2\varepsilon\, x(t)\right] \qquad \text{(VI-32)}$$

or

$$dA^*\lfloor x,1 \rfloor = R \left\{ dA\lfloor x,1 \rfloor \right\} + 2\varepsilon \lfloor R \left\{ x(t) \right\} - x(t) \rfloor \qquad \text{(VI-33)}$$

The above expression for $dA^*\lfloor x,1 \rfloor$ is used in the computer program.

The second differential of the objective function, or the first differential of the gradient of the performance index is denoted by $g(x)$. It is derived from Equation (VI-18) after substituting for $dA^*\lfloor x,1 \rfloor$ the expression given by Equation (VI-23). Thus:

$$g(x) = 2\lambda + 2\varepsilon \qquad \text{(VI-34)}$$

The step-size $1/g(x)$ remains fixed throughout the search procedure. Newton's iteration formula for the constrained problem is written as:

$$x_{n+1} = x_n - \frac{2\lambda x_n - dA^*\lfloor x_n, 1 \rfloor}{2(\lambda + \varepsilon)} \qquad \text{(VI-35)}$$

Whenever the control signal lies within the constraint boundary we set $\lambda = 0$ in Equation (VI-35).

In this problem the number of samples in the interval (0,T) was varied for two computer runs. In one case it was set equal to 4 and in another equal to 21. The change reflects on the accuracy of the integration subroutines.

With a step-size equal to $1/g(x)$ the algorithm converges in one step for $\lambda = 0$. Depending on the density of the grid sample points subdividing the interval (0,T) the control signal is given very accurately by Figure (VI-8). With 21 samples and a step-size reduced by a factor of four, the scheme converges within ten iteration steps. With a severe input constraint (p = 2 in Equation VI-24) and a reduced step-size, the control signal reaches the boundary in two steps.

When a magnitude constraint is placed on the input function the adaptive procedure converges in four steps as shown in Figure (VI-9) (a) and (b). In all cases the arbitrary initial value for x(t) in conjunction with the magnitude of the step-size was found to be very critical for the convergence of the search technique.

The computer programs for the case presented above are included in Appendix II.

VI - 3. Case Study Number 3.

We consider the operation of a single, continuous stirred-tank reactor [VI-1, 2] with a first-order exothermic reaction (Figure VI-10). The reaction is: $A \rightarrow B$.

If c is the concentration of A in the product and on the assumption that a cooling coil is used to remove the heat generated by the chemical reaction, the dynamic or transient mass balance is given by:

$$V \frac{dc}{dt} = Fc_o - Fc - Vke^{-E/RT} c \qquad \text{(VI-36)}$$

and the corresponding heat balance by[+] :

$$V \rho c_p \frac{dT}{dt} = F \rho c_p (T_o - T) - \Delta HVke^{-E/RT} c - U(T-T') \qquad \text{(VI-37)}$$

where:

- V = reaction volune
- F = feed flow rate
- k = reaction constant
- E = energy of activation
- ρ = density of solution
- c_p = heat capacity of solution
- H = heat of reaction
- U = heat transfer coefficient for heating coil

[+] For the derivation of these equations see References VI-1 and VI-2.

T' = average cooling-water temperature

T_o = initial temperature in reactor

c_o = initial concentration in reactor

If y is the output concentration then no mass balance equation is needed for B since

$$y - y_o = c_o - c \tag{VI-38}$$

and knowing c, we can calculate y directly from Equation (VI-38).

It is ordinarily assumed that ρ, F and V are constant and perfect mixing occurs i.e. c is the same in the reactor and product stream. In reality these parameters and the reaction constant k vary with the operation of the reactor. It is therefore possible to view the control of the stirred-tank reactor as an adaptive problem.

Equations (VI-36) and (VI-37) can be simplified by defining:

$$\alpha_1 = \frac{Fc_o}{V} \qquad \alpha_2 = \frac{FT_o}{V} + \frac{UT'}{V\rho c_p}$$

$$\beta_1 = \frac{F}{V} \qquad \beta_2 = \frac{F}{V} + \frac{U}{V\rho c_p}$$

$$\gamma_1 = k \qquad \gamma_2 = \frac{(-\Delta H)k}{\rho c_p}$$

which leads to:

$$\frac{dc}{dt} = \alpha_1 - c(\beta_1 + \gamma_1 e^{-E/RT}) \quad \text{(VI-39)}$$

$$\frac{dT}{dt} = \alpha_2 - \beta_2 T + \gamma_2 c e^{-E/RT} \quad \text{(VI-40)}$$

The system is initial value and consists of two first-order coupled nonlinear ordinary differential equations. The reactor temperature is taken as the control variable. It is assumed that proper monitoring and temperatare transducing instrumentation is available.

We seek to maximize the output product y in a time interval (0,T) assuming, of course, stable operation. A magnitude constraint is placed on the reactor temperature which is designated from here on by x, reserving the letter T for the optimization interval. The augmented performance functional is therefore given by:

$$P[x] = - \int_0^T y dt + \lambda \max |x| \quad \text{(VI-41)}$$

and a minimum of $P\lfloor x \rfloor$ is sought. The solution leads to the following expression for the gradient of $P\lfloor x \rfloor$:

$$\nabla P\lfloor x \rfloor = x - \max |x| \operatorname{sgn} dy^*\lfloor x, 1 \rfloor \quad \text{(VI-42)}$$

For the identification of the adjoint we consider Equation (VI-34) in normalized form; with $y = c_o - c$ it may be written as:

$$\frac{dy}{dt} = -1 + (c_o - y)\ (1 + e^{-\frac{1}{x}}) \qquad \text{(VI-43)}$$

By definition a measure of the output variation $\delta y(x, \delta x)$, when the input is perturbed by δx, is given by:

$$\delta y(x, \delta x) = y(x + \delta x) - y(x) \qquad \text{(VI-44)}$$

Introducing the output variation δy and the input variation δx into Equation (VI-43) and keeping only first order terms results in

$$\frac{d}{dt}\ \delta y = F(t)\ \delta y + H(t)\ \delta x \qquad \text{(VI-45)}$$

With a particular variation $\overline{\delta x}$ equal to unity the adjoint differential is approximately given by:

$$dy^*[x, 1] = R\ \left\{ \delta y(x, \overline{\delta x})/\gamma \right\} \qquad \text{(VI-46)}$$

with $\overline{\delta x} = \gamma h$ and assuming that $F(t)$ and $H(t)$ remain fairly constant in the interval (0,T). The values of F(t) and H(t) are continuously monitored in a computer simulation; mathematically, they are expressed by:

$$F(t) = \frac{\partial}{\partial y} \left(\frac{dy}{dt} \right) = -1-e^{-\frac{1}{x}} \tag{VI-47}$$

$$H(t) = \frac{\partial}{\partial x} \left(\frac{dy}{dt} \right) = (c_o - y) \frac{1}{x^2} e^{-\frac{1}{x}} \tag{VI-48}$$

Starting with two arbitrary initial points for the concentration c and the reactor temperature x a fourth-order Runge-Kutta method, properly expanded to handle two dependent variables, is used to calculate the transient response of the reactor [VI-3]. The step-size is set equal to 0.5 and six values for x(t) and y(t) are computed. This transient solution for x(t), $t \in (0,T)$, is used as the arbitrary initial input $x_o(t)$ that is necessary for the estimation of the immediately following element of the input sequence, utilizing the adaptive control loop. Thus, Equation (VI-40) is treated as a differential constraint on the input function and the minimization of $P\lfloor x \rfloor$ proceeds under the assumption that this constraint is satisfied always.

A listing of the computer program is included in Appendix II. The plant differentials were calculated with the parameter γ set equal to 0.1 A fixed step-size of magnitude equal to 0,1 was used in the iteration algorithm. The input function reached the boundary of the constraint in 78 iteration steps and the execution time was 1.95 minutes. The results are found to be satisfactory only in the sense that the output product tends towards a maximum. The example illustrates the applicability of the basic adaptive algorithm without claiming physical consistency for the obtained results.

Figure (VI-11) shows the output and control functions vs. time. The multipliers F(t) and H(t) did not vary appreciably in the interval (0,T) making any subdivision of the identification period unnecessary.

VI - 4. A Comparison.

It is demonstrated below that for the example considered in Section VI-1 the integral operator approach leads to identical results with those suggested by a differential operator formulation.

Consider Equation (VI-3) which is repeated here for convenience:

$$y(t) = y_o + \int_0^t c' e^{-b(t-\tau)} [1-e^{-ax(\tau)}] d\tau \qquad \text{(VI-49)}$$

Applying Leibnitz's rule to Equation (VI-49) results in:

$$\frac{dy}{dt} = \left[1-e^{-ax(t)}\right]\cdot 1 + (-b)\int_0^t c'e^{-b(t-\tau)}\left[1-e^{-ax(\tau)}\right]d\tau \tag{VI-50}$$

or

$$\frac{dy}{dt} = 1-e^{-ax(t)} - b\left[y(t)-y_o\right] \tag{VI-51}$$

Equation (VI-51) is recognized to be of the form:

$$\frac{dy}{dt} = f(x,y,c) \tag{VI-52}$$

and therefore the identification and optimization procedures introduced in Chapter IV for plants described by nonlinear differential equations apply to Equation (VI-51). Following the standard procedure we introduce the variations δy and δx and write the following linearized version of Equation (VI-51):

$$\frac{d}{dt}\,\delta y = F(t)\,\delta y + H(t)\,\delta x \tag{VI-53}$$

with

$$\left.\begin{aligned} \delta y &= y_{n+1} - y_n \\ \delta x &= x_{n+1} - x_n \end{aligned}\right\} \qquad \text{(VI-54)}$$

and

$$\left.\begin{aligned} F(t) &= -b \\ H(t) &= ae^{-ax(t)} \end{aligned}\right|_{\begin{aligned} x(t) &= x_n \\ y(t) &= y_n \end{aligned}} \qquad \text{(VI-55)}$$

The output variation $\delta y(t)$ is found to be equal to:

$$\delta y(t) = \Phi(t,0)\,\delta y(0) + \int_0^t \Phi(t,\tau)H(\tau)\,\delta x(\tau)d\tau \qquad \text{(VI-56)}$$

And the adjoint differential $B^*g(y)$ is given by Equation (IV-34), or

$$B^*g(y) = H(t)\int_t^T \Phi(\tau,t)g[y(\tau)]d\tau \qquad \text{(VI-57)}$$

The adjoint differential for the plant described by the integral equation (VI-49) is written as

$$dA^*[x,g] = ae^{-ax(t)} \int_t^{\tau} e^{-b(\tau-t)} g(\tau)d\tau \qquad \text{(VI-58)}$$

It is true that $H(t) = ae^{-ax(t)}$ and it can be easily shown that for the linear system of Equation (VI-53):

$$\Phi(\tau,t) = e^{-b(\tau-t)} \qquad \text{(VI-59)}$$

Identical expressions for the adjoint differential have been established when the input-output relationship for the same plant is of the integral or differential type.

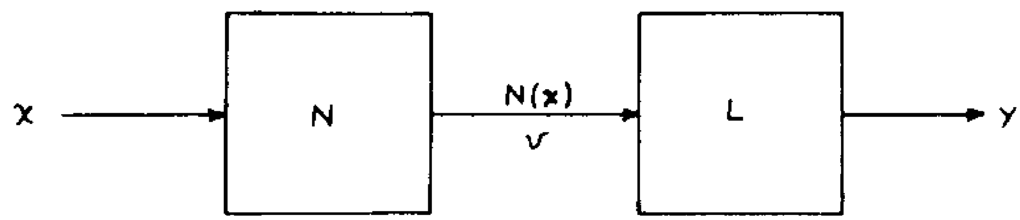


FIGURE VI - 1

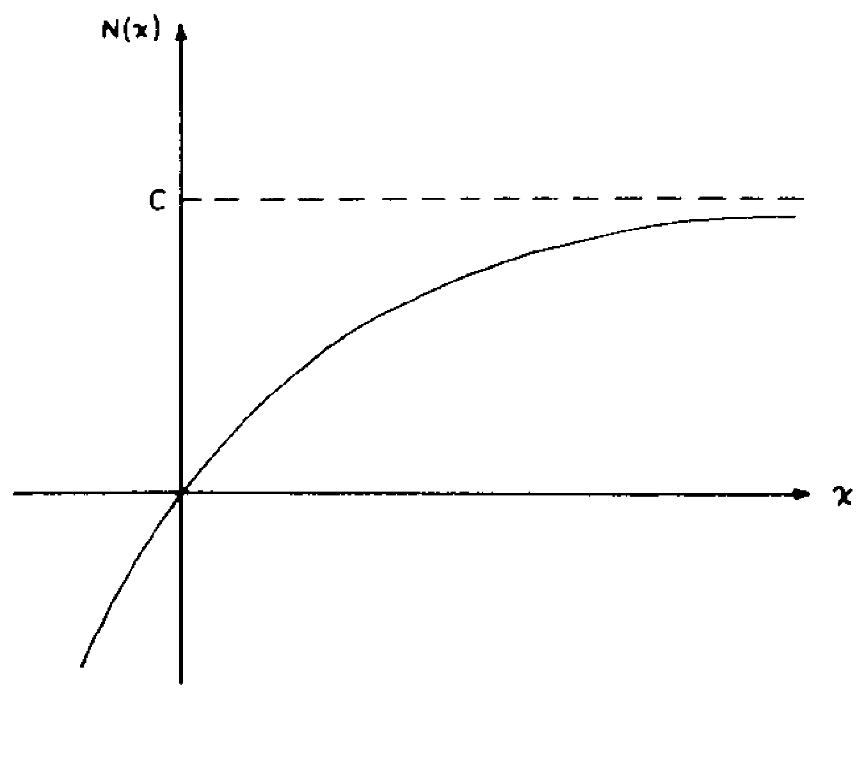


FIGURE VI - 2

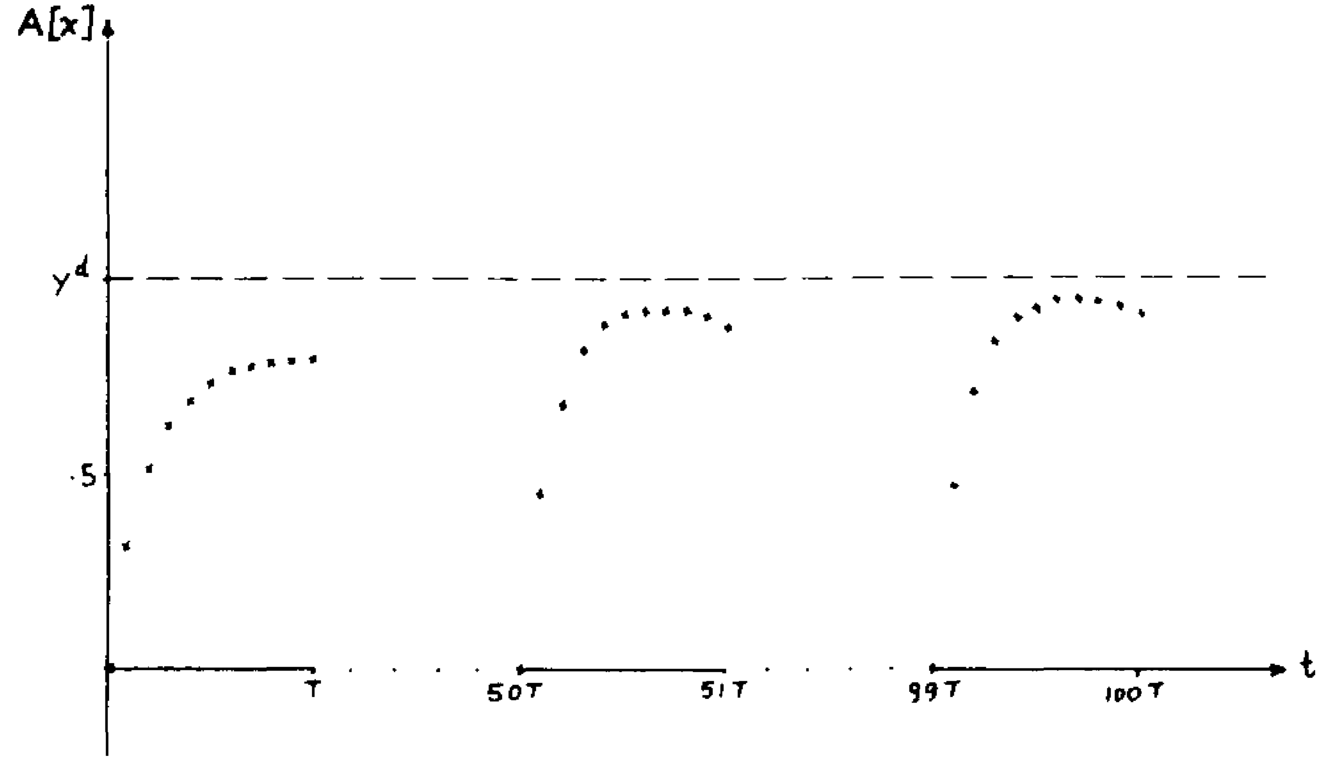


FIGURE VI - 3 (a)

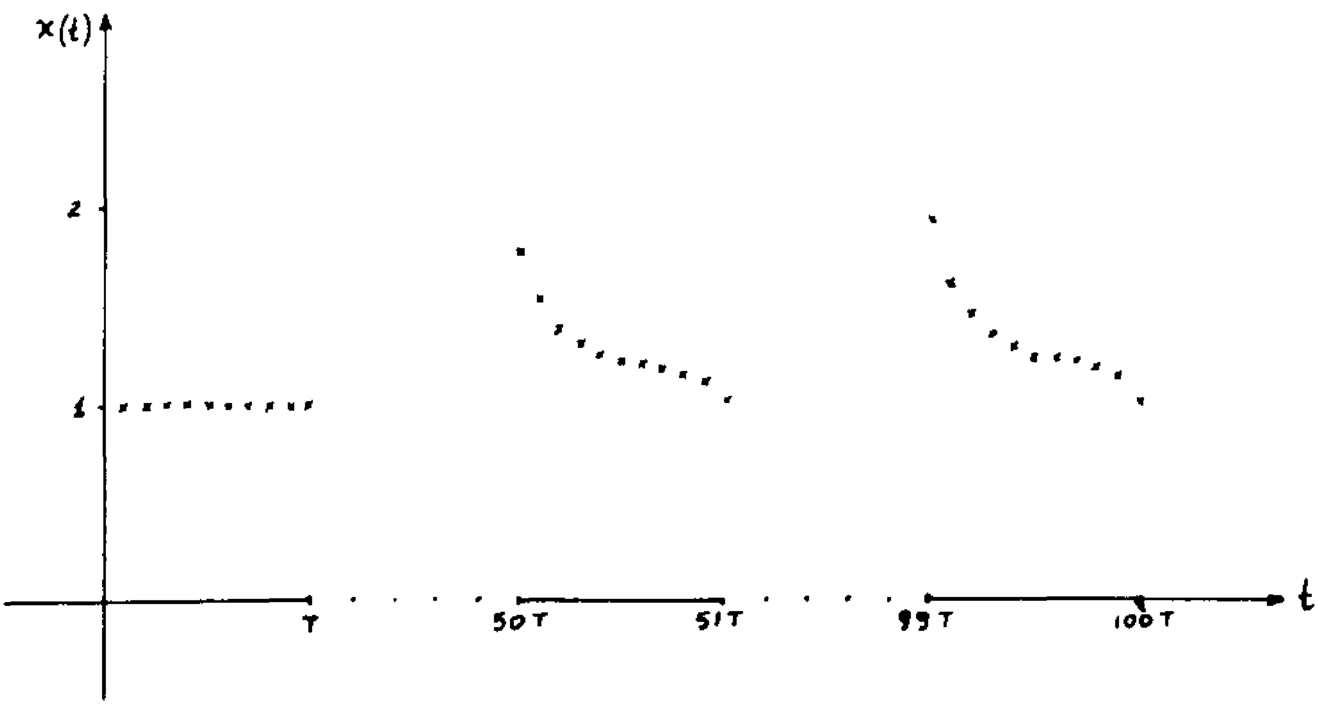


FIGURE VI - 3 (b)

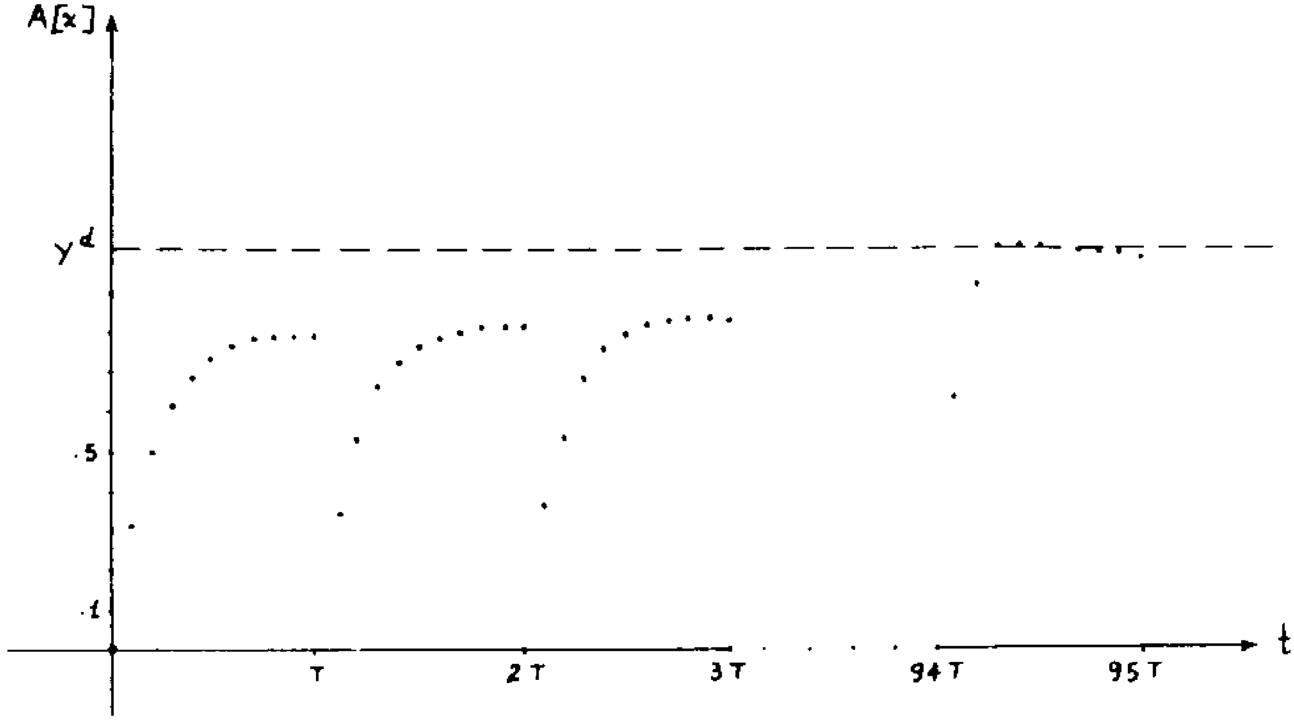


FIGURE VI-4 (a)

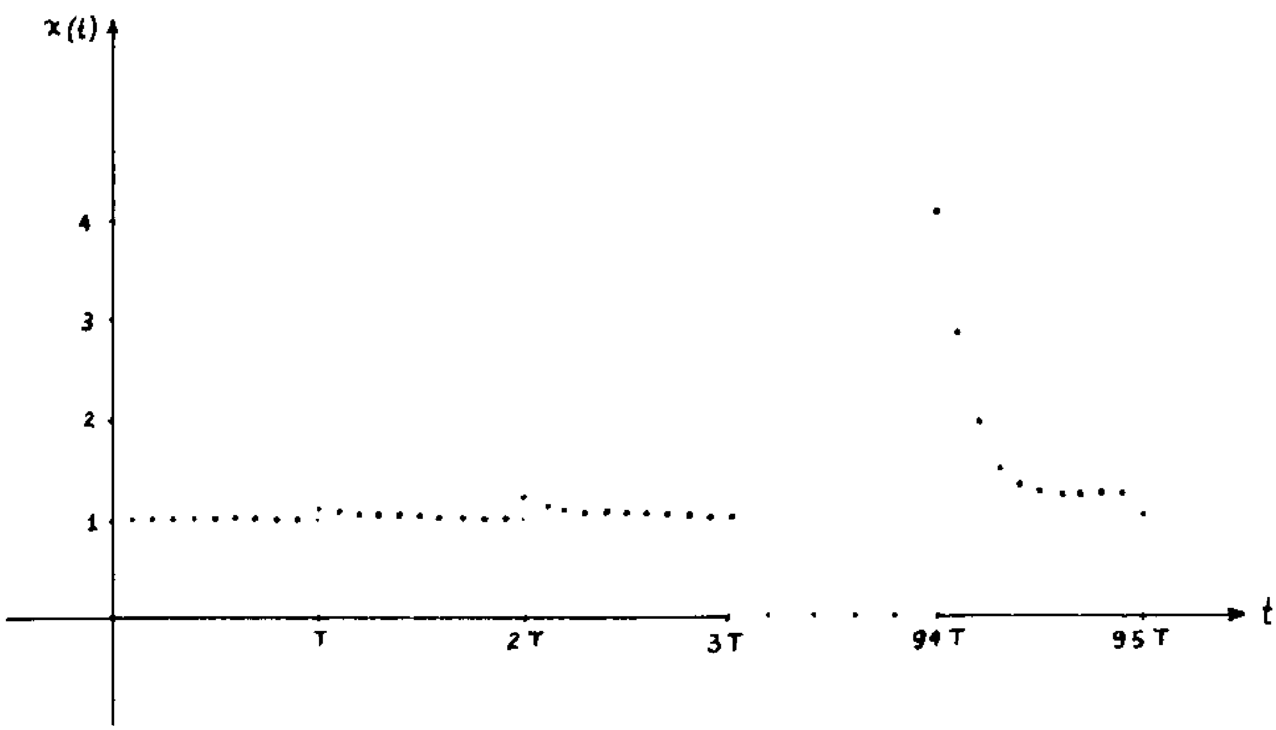


FIGURE VI-4 (b)

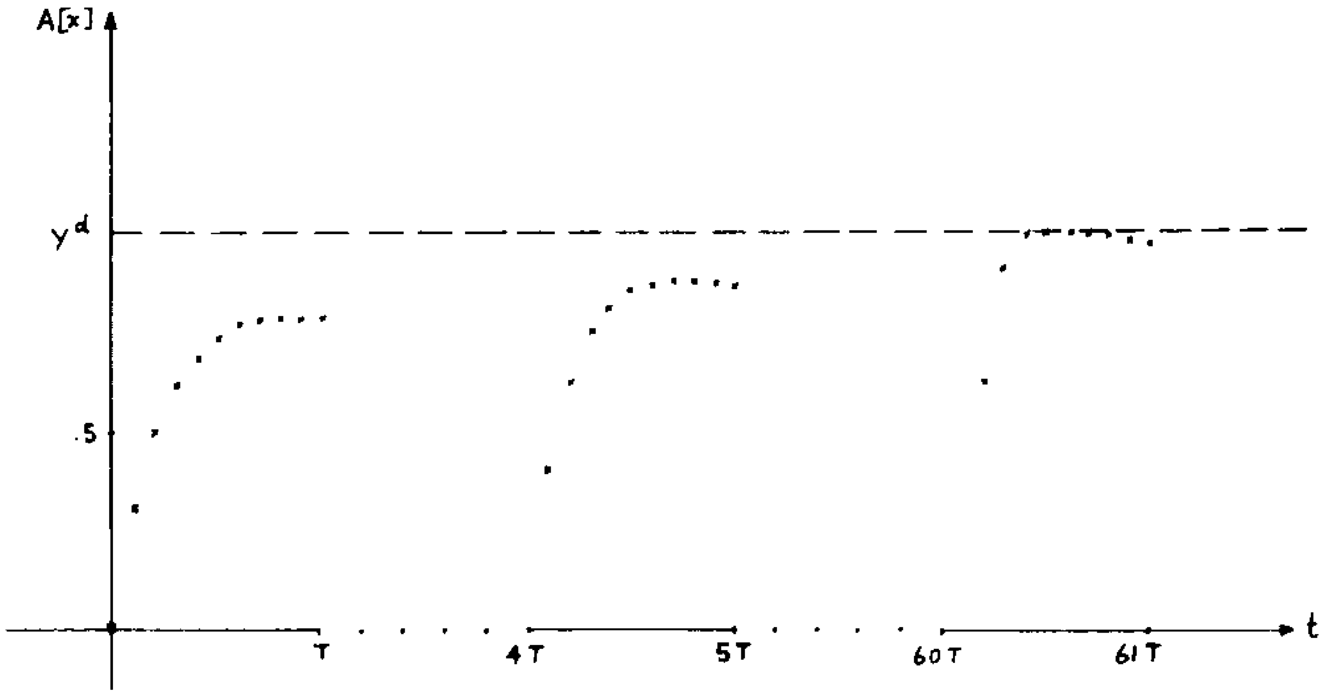


FIGURE VI - 5 (a)

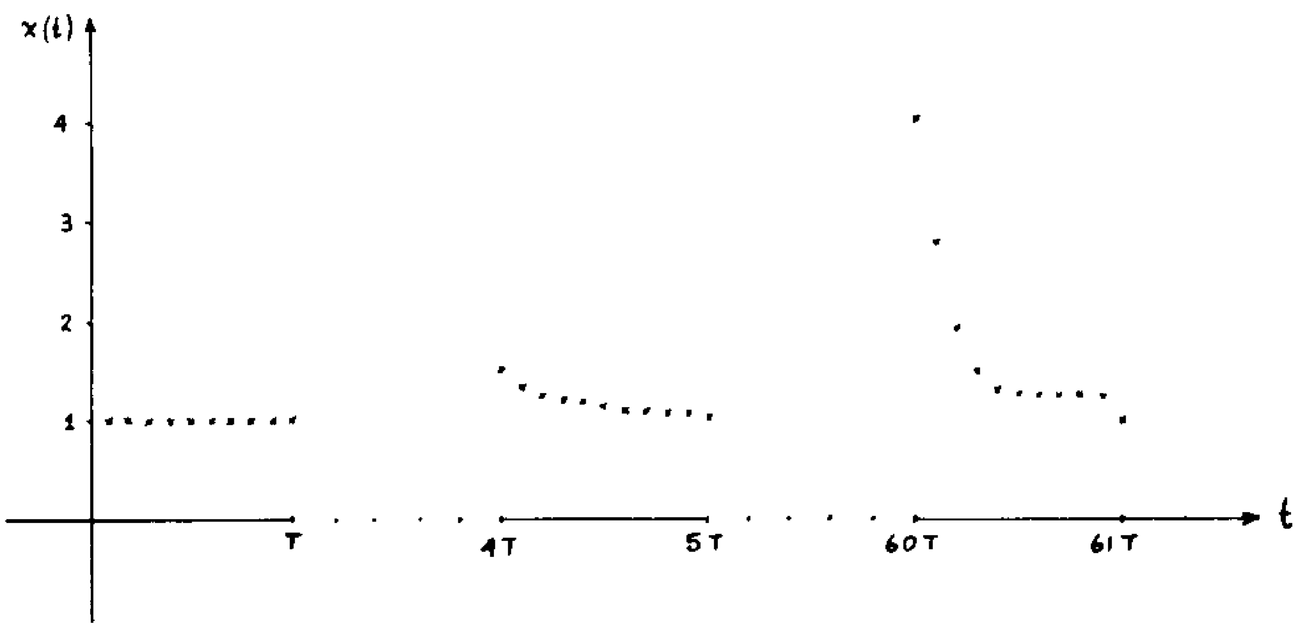


FIGURE VI - 5 (b)

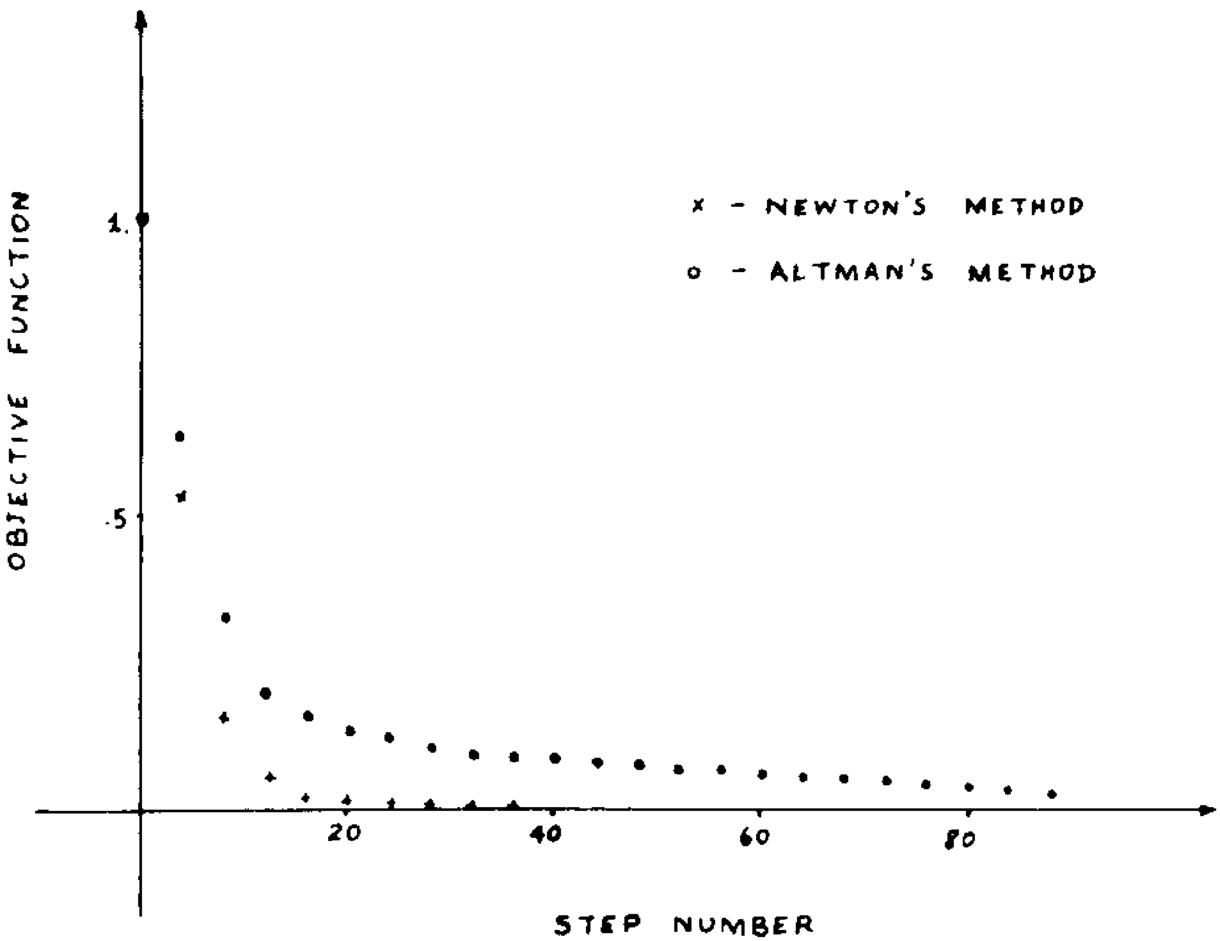

OBJECTIVE FUNCTION
1.
.5
x - NEWTON'S METHOD
o - ALTMAN'S METHOD
20
40
60
80
STEP NUMBER

FIGURE VI - 6

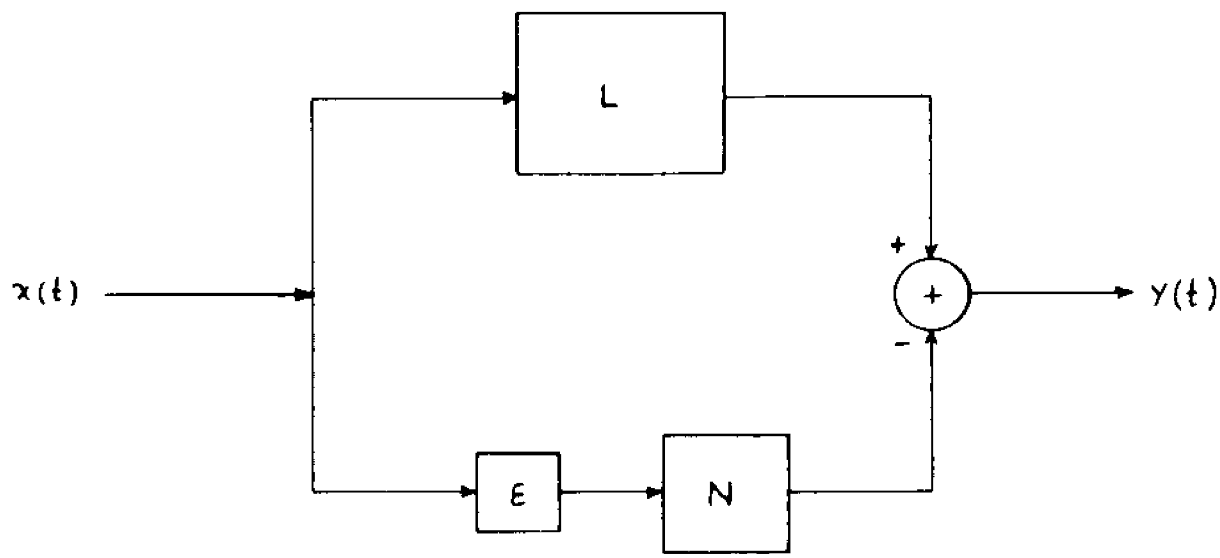


FIGURE VI - 7

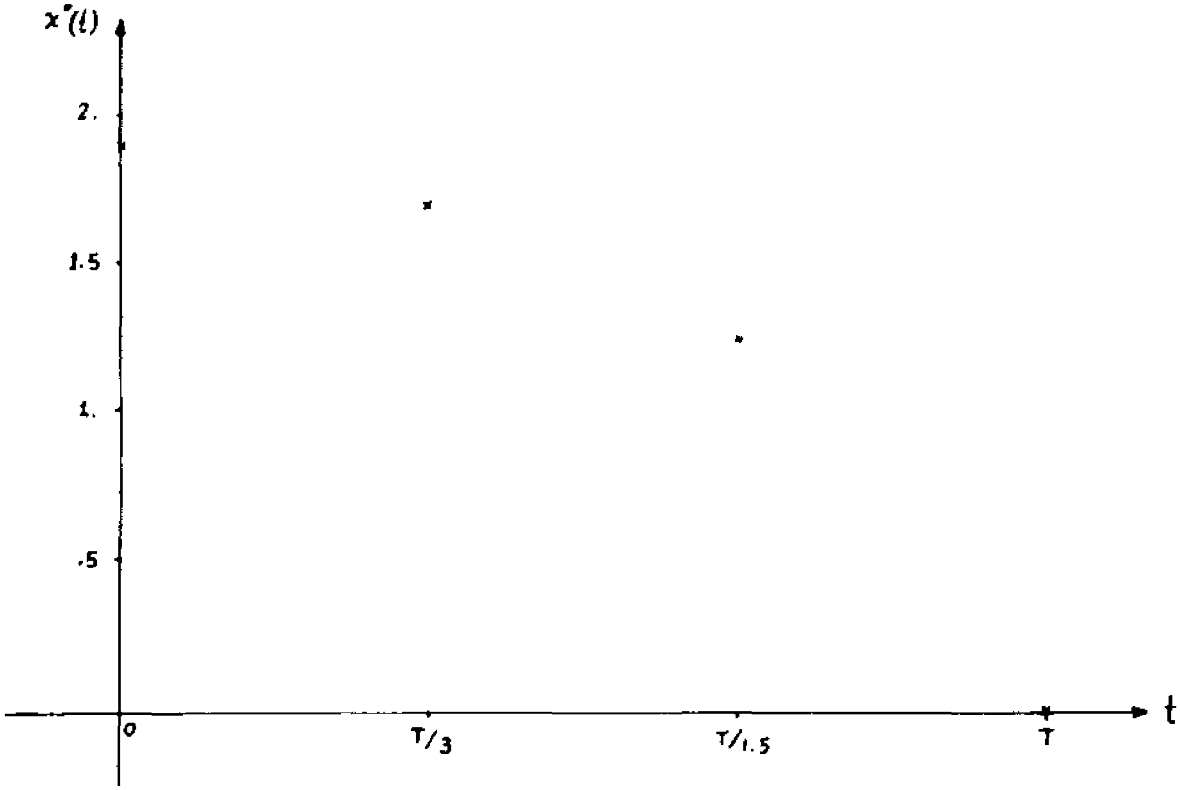


FIGURE VI - 8

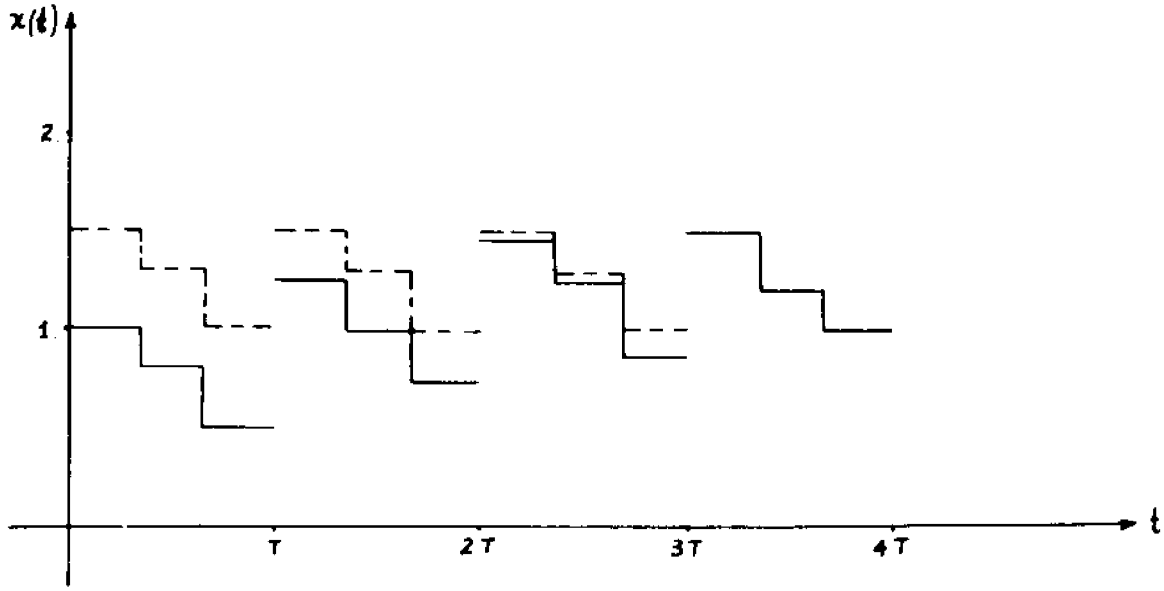


FIGURE VI-9 (a)

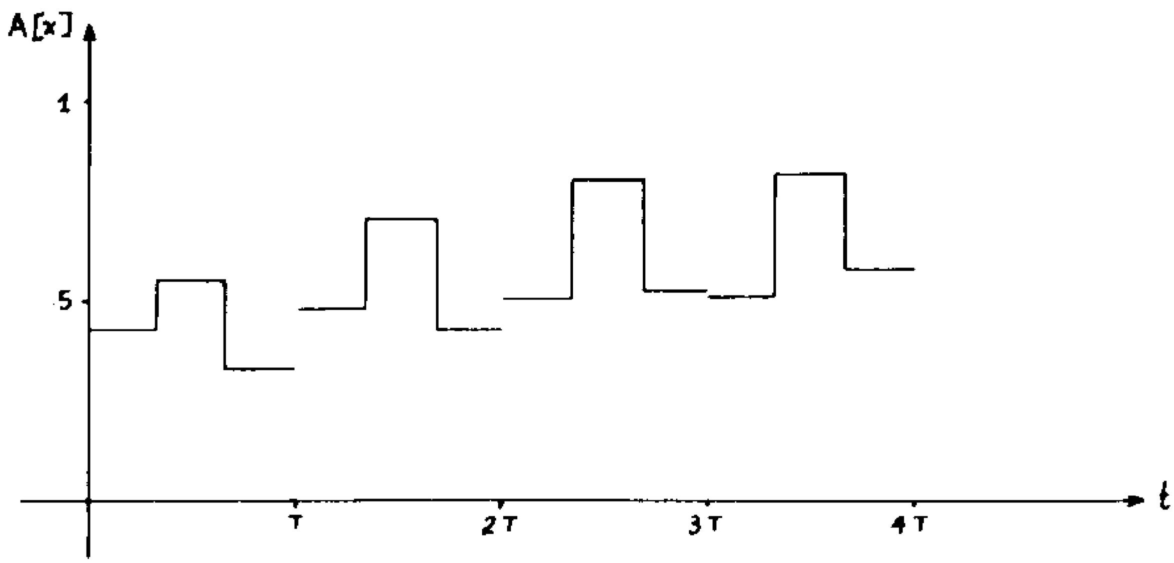


FIGURE VI-9 (b)

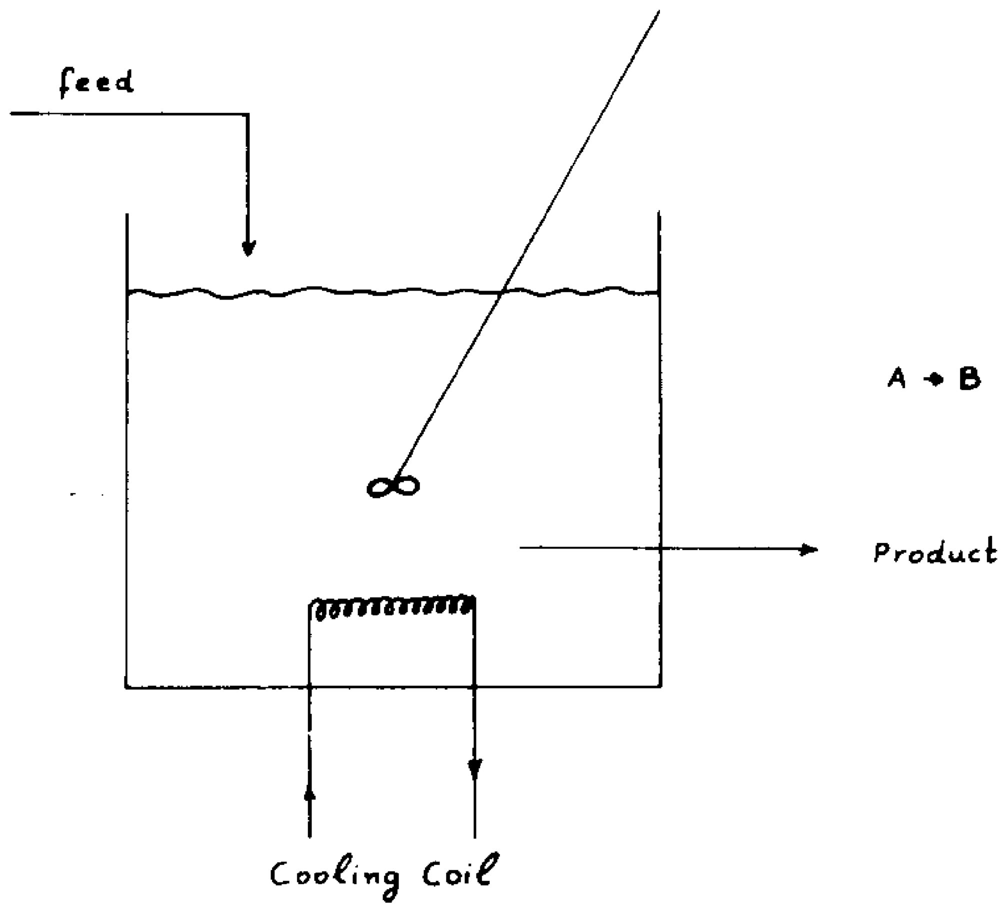


FIGURE VI - 10

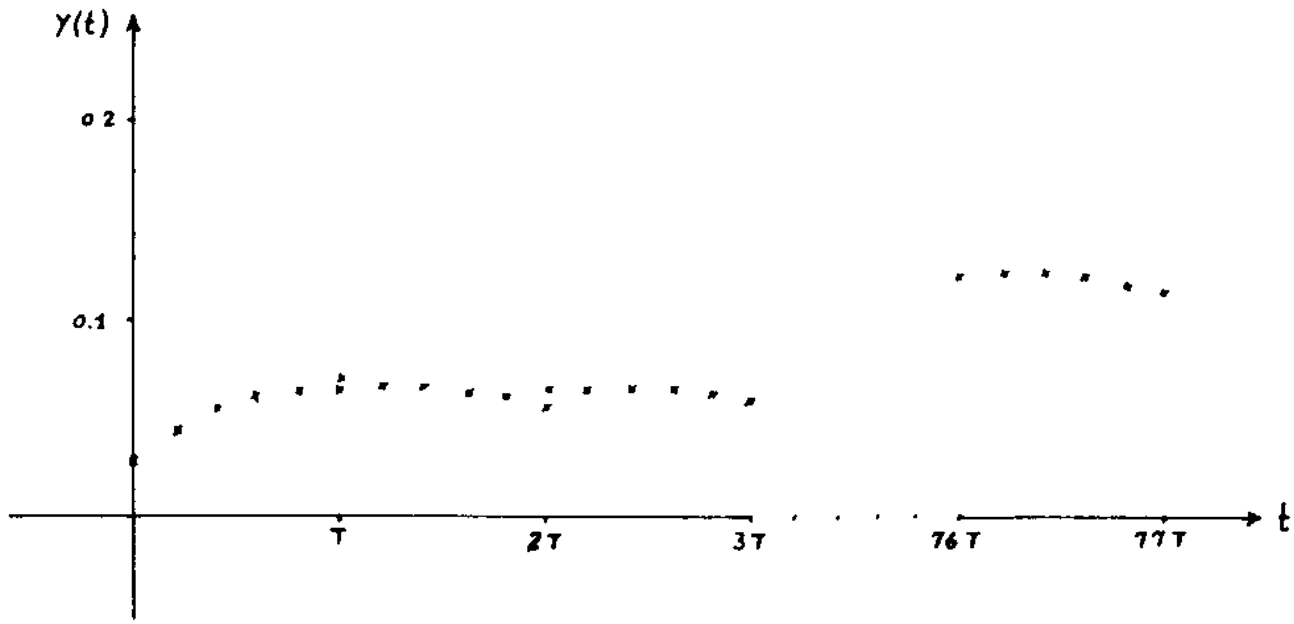


FIGURE VI-11 (a).

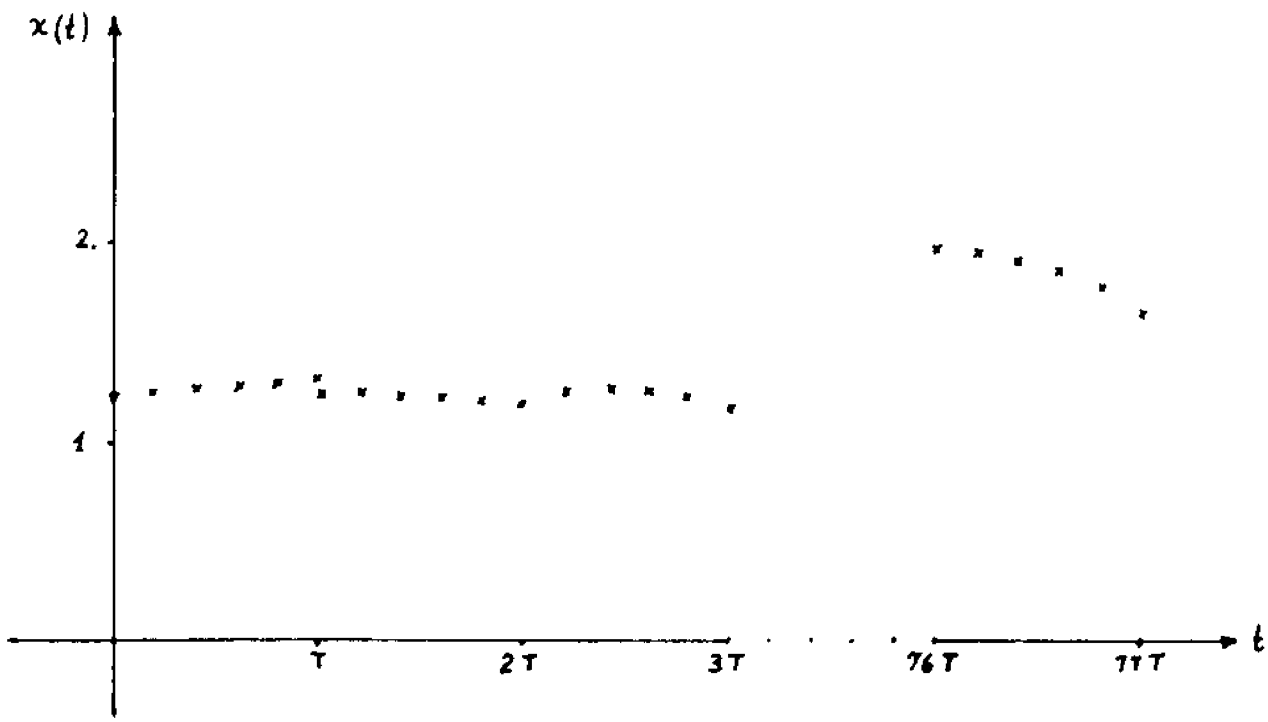
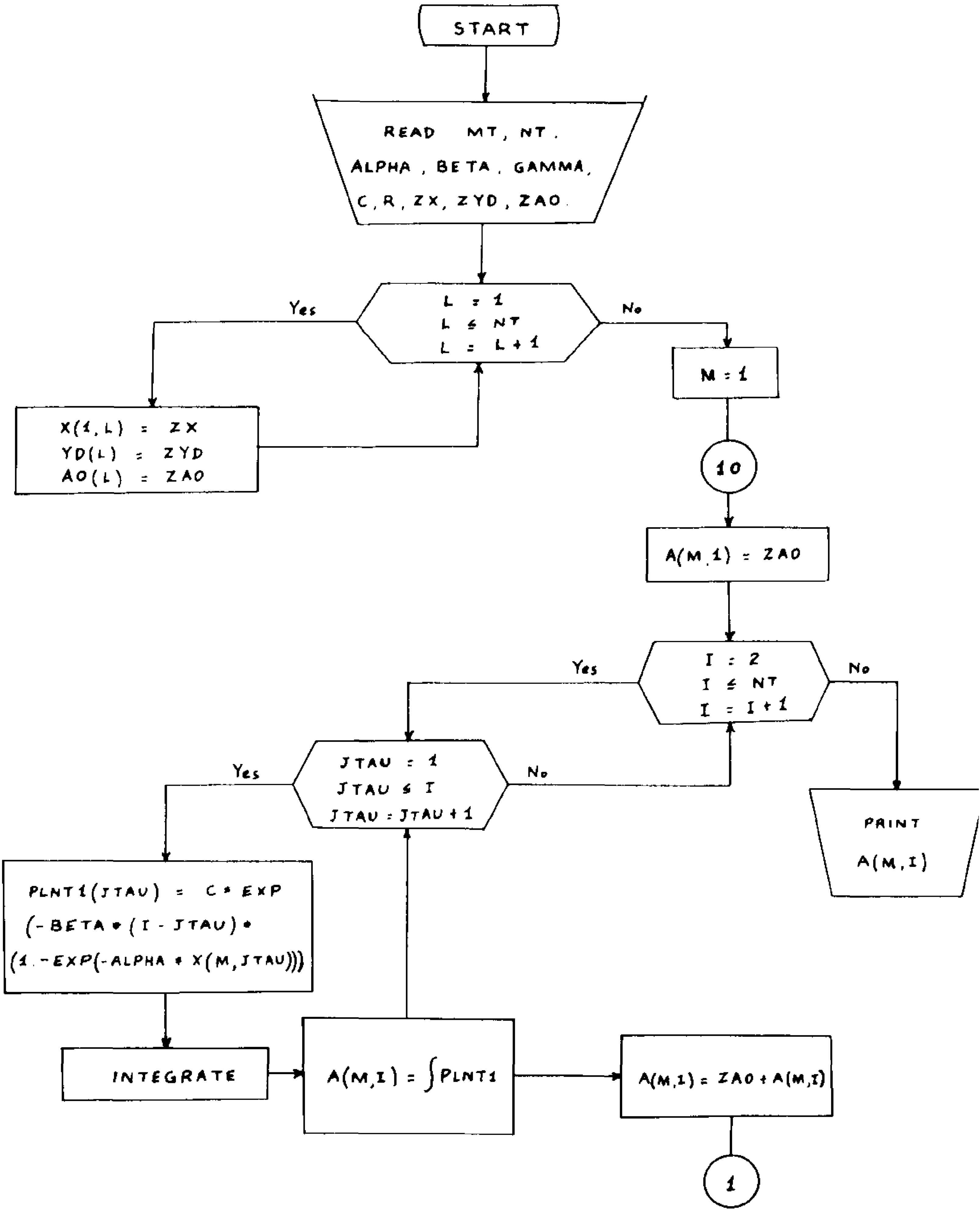
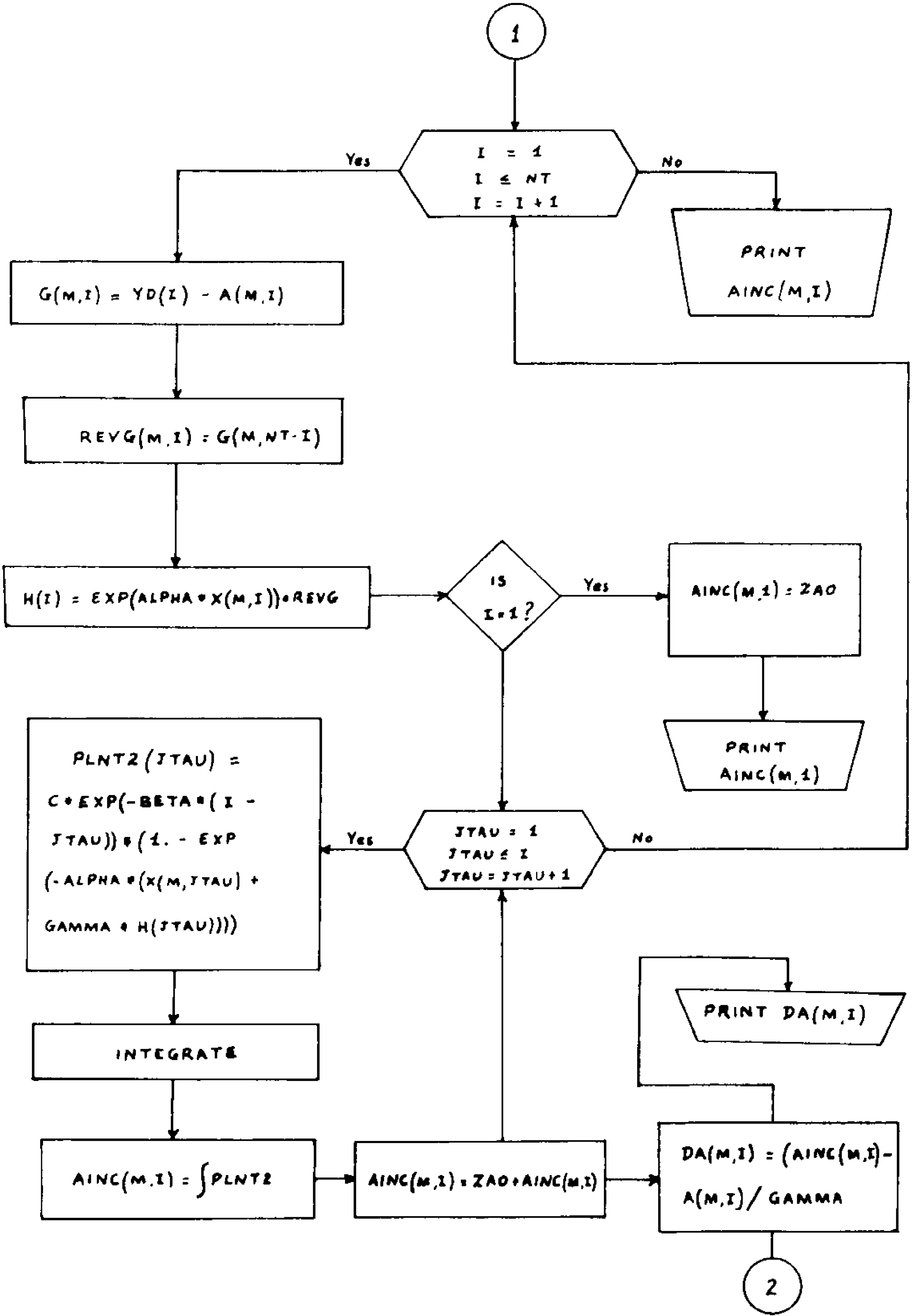
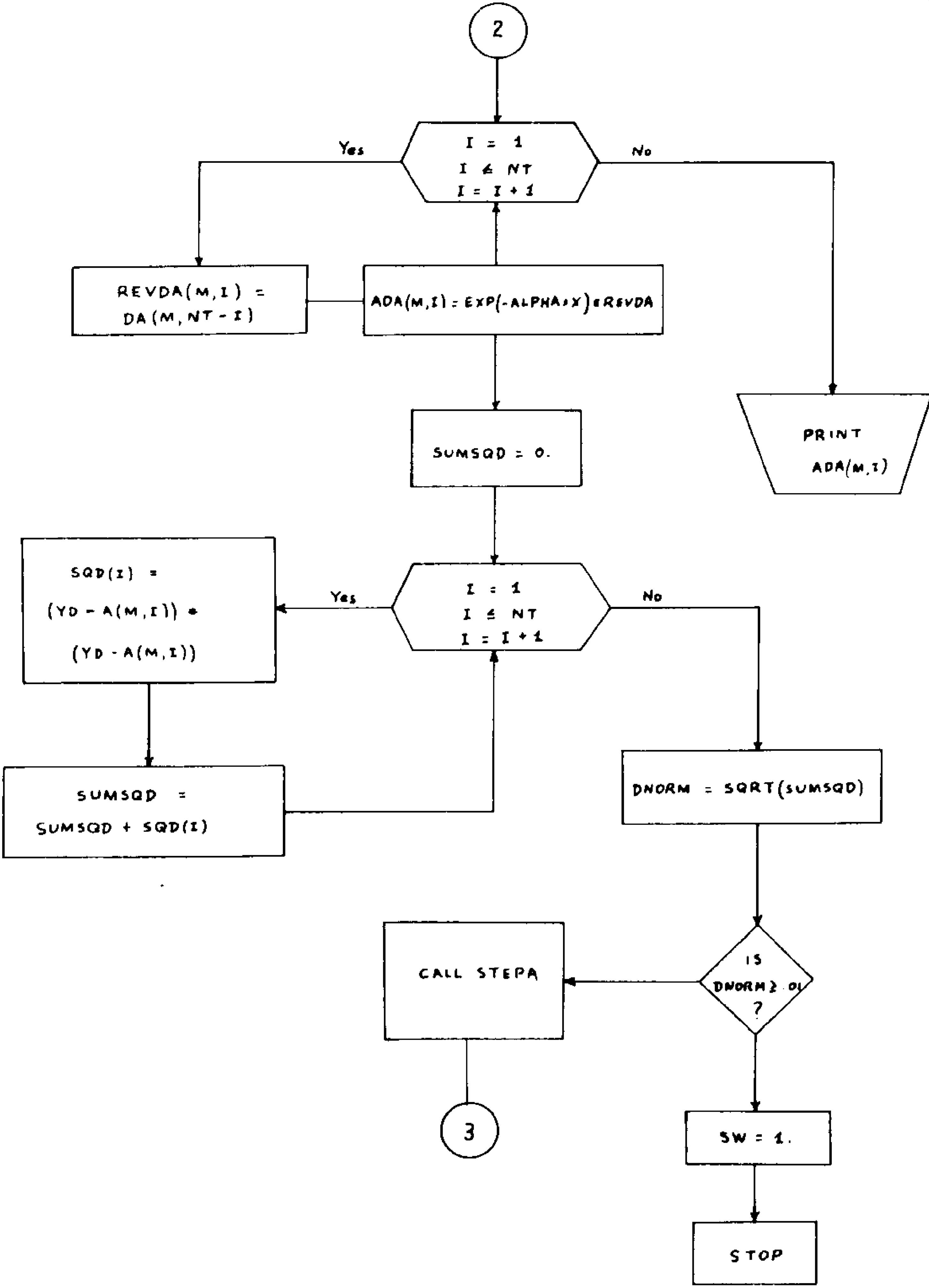
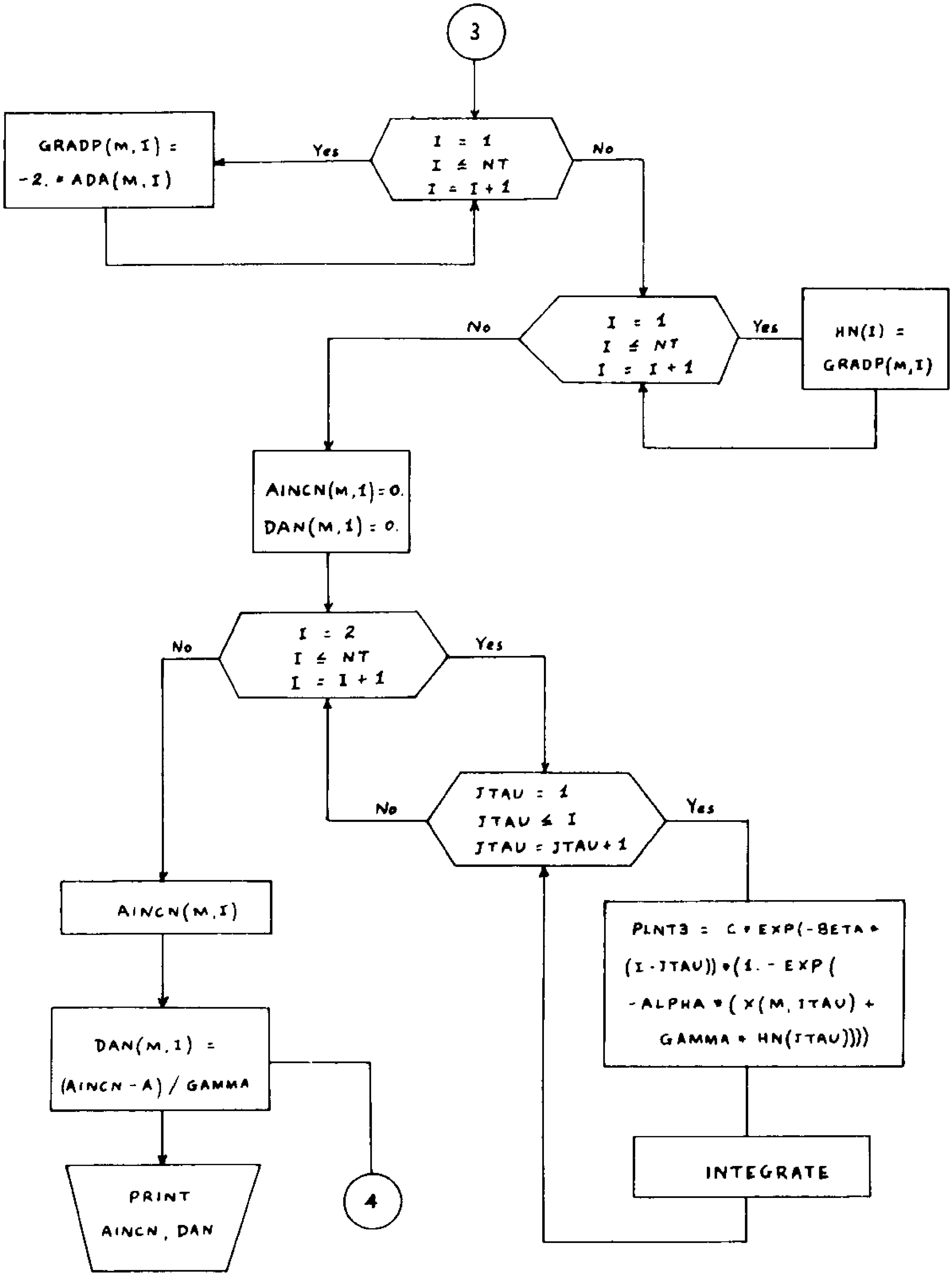


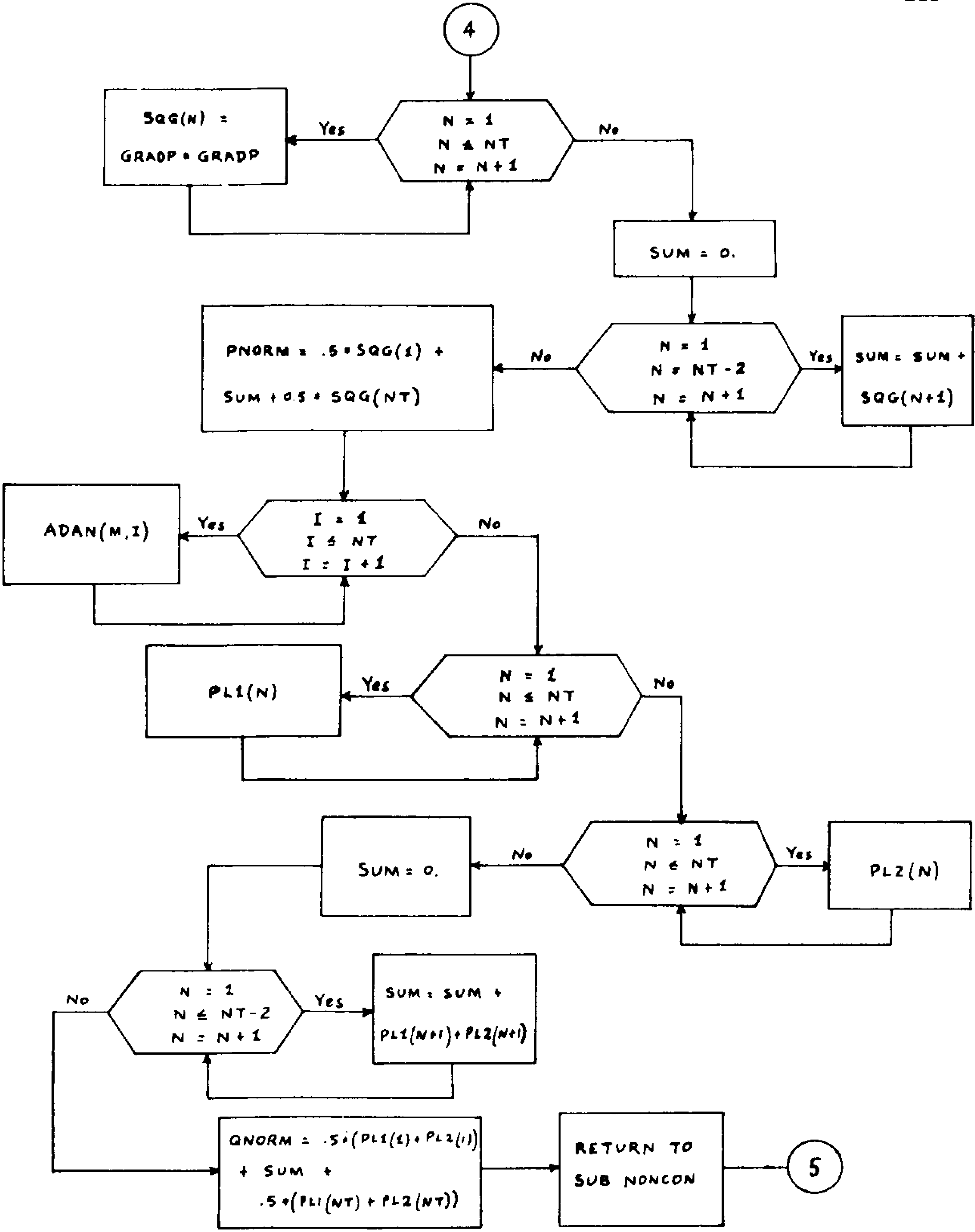
FIGURE VI-11 (b).

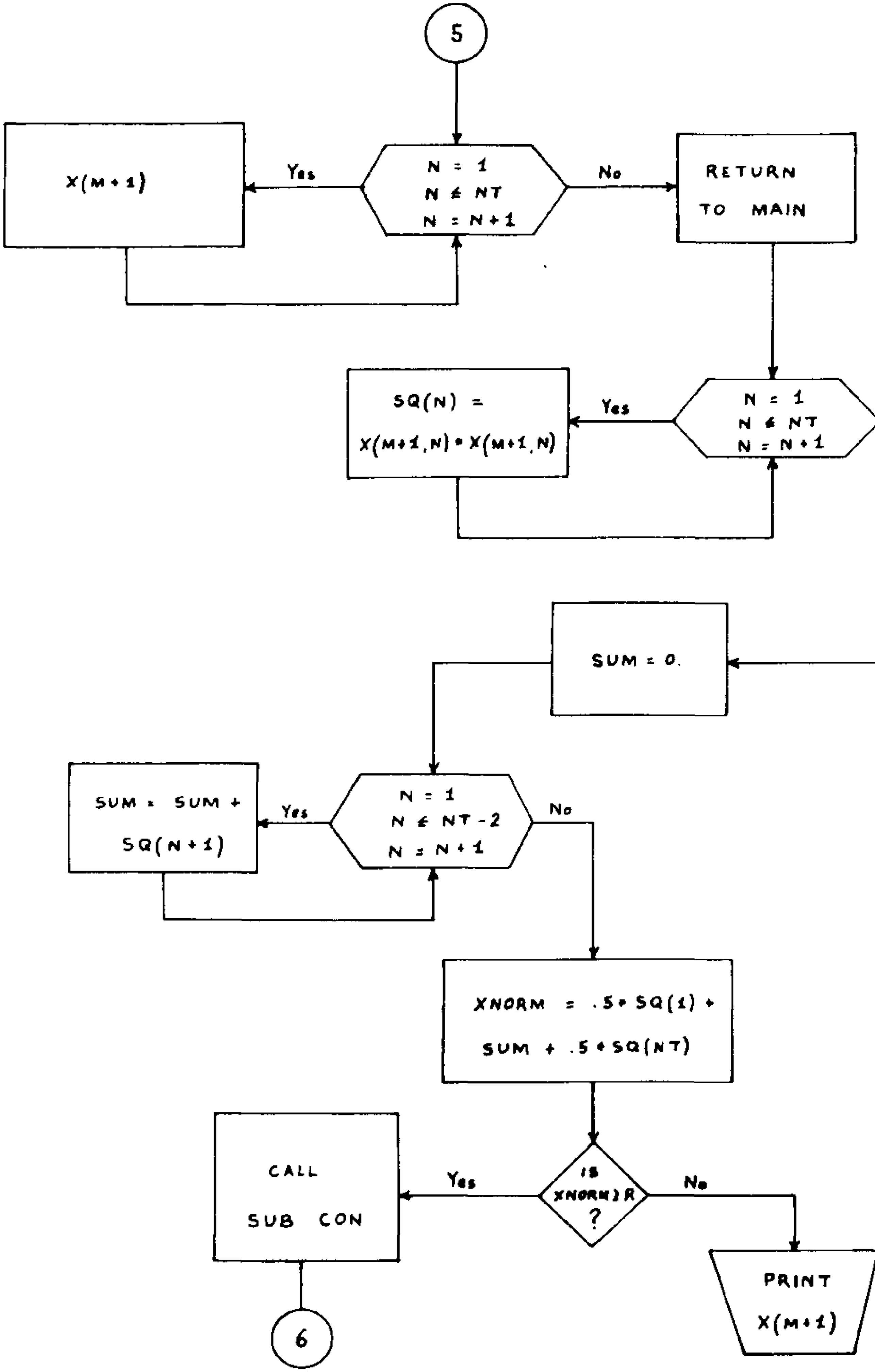

START
READ MT, NT.
ALPHA, BETA, GAMMA,
C, R, ZX, ZYD, ZAO.
L = 1
L ≤ NT
L = L + 1
Yes
No
X(1, L) = ZX
YD(L) = ZYD
AO(L) = ZAO
M = 1
10
A(M, 1) = ZAO
I = 2
I ≤ NT
I = I + 1
Yes
No
JTAU = 1
JTAU ≤ I
JTAU = JTAU + 1
Yes
No
PRINT
A(M, I)
PLNT1(JTAU) = C * EXP
(-BETA * (I - JTAU) *
(1. - EXP(-ALPHA * X(M, JTAU)))
INTEGRATE
A(M, I) = ∫PLNT1
A(M, 1) = ZAO + A(M, I)
1

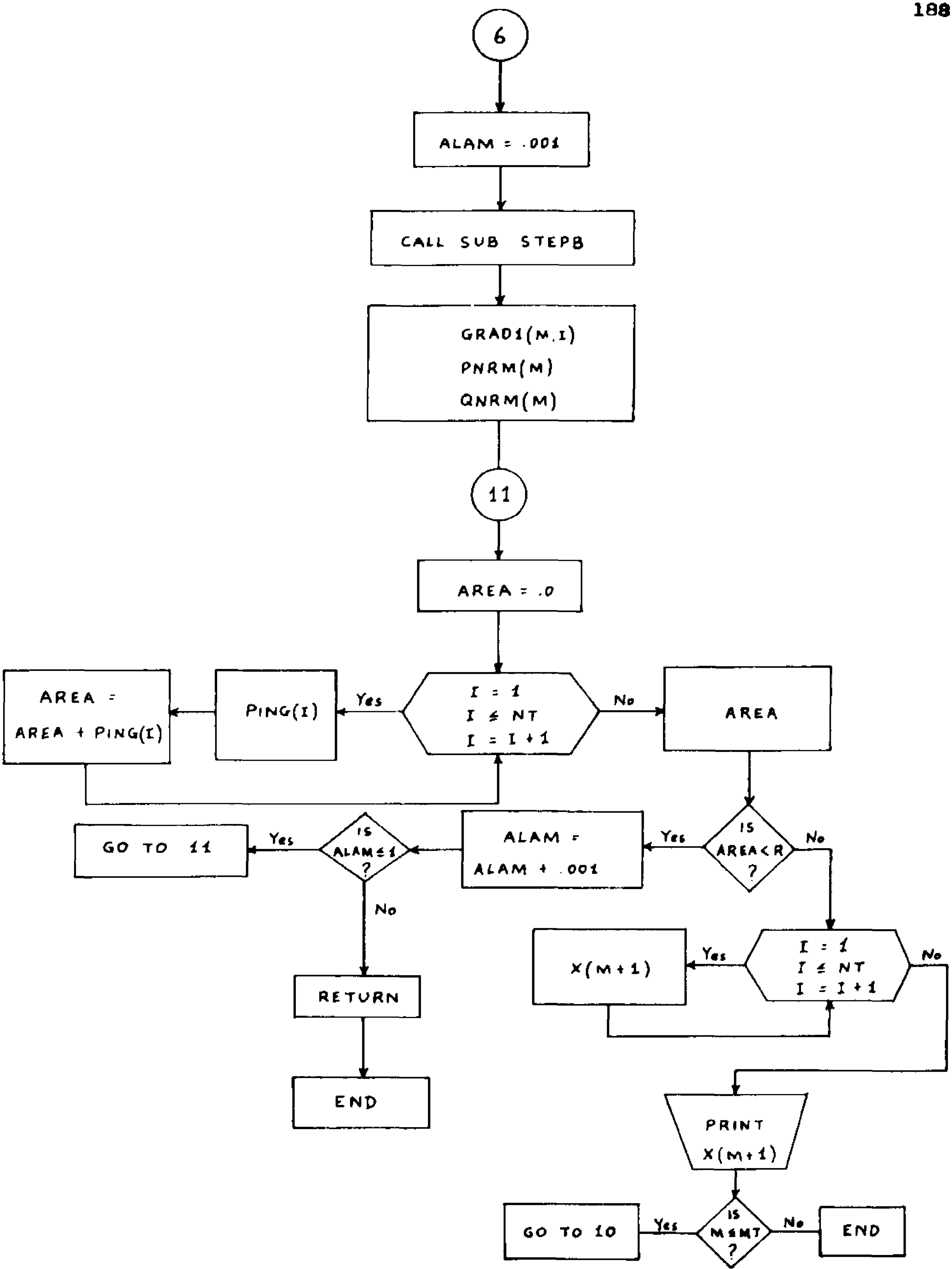

1
I = 1
I ≤ NT
I = I + 1
Yes
No
PRINT
AINC(M,I)
G(M,I) = YD(I) - A(M,I)
REVG(M,I) = G(M,NT-I)
H(I) = EXP(ALPHA * X(M,I)) * REVG
IS
I = 1 ?
Yes
AINC(M,1) = ZAO
PRINT
AINC(M,1)
PLNT2(JTAU) =
C * EXP(-BETA * (I -
JTAU)) * (1. - EXP
(-ALPHA * (X(M,JTAU) +
GAMMA * H(JTAU))))
Yes
JTAU = 1
JTAU ≤ I
JTAU = JTAU + 1
No
PRINT DA(M,I)
INTEGRATE
AINC(M,I) = ∫PLNT2
AINC(M,I) = ZAO + AINC(M,I)
DA(M,I) = (AINC(M,I) -
A(M,I) / GAMMA
2


2
I = 1
I ≤ NT
I = I + 1
Yes
No
REVDA(M,I) =
DA(M,NT - I)
ADA(M,I) = EXP(-ALPHA*X)*REVDA
PRINT
ADA(M,I)
SUMSQD = 0.
I = 1
I ≤ NT
I = I + 1
Yes
No
SQD(I) =
(YD - A(M,I)) *
(YD - A(M,I))
SUMSQD =
SUMSQD + SQD(I)
DNORM = SQRT(SUMSQD)
IS
DNORM ≥ .01
?
CALL STEPA
3
SW = 1.
STOP


3
I = 1
I ≤ NT
I = I + 1
Yes
No
GRADP(M,I) =
-2. * ADA(M,I)
I = 1
I ≤ NT
I = I + 1
No
Yes
HN(I) =
GRADP(M,I)
AINCN(M,1) = 0.
DAN(M,1) = 0.
I = 2
I ≤ NT
I = I + 1
No
Yes
JTAU = 1
JTAU ≤ I
JTAU = JTAU + 1
No
Yes
PLNT3 = C * EXP(-BETA *
(I - JTAU)) * (1. - EXP(
-ALPHA * (X(M, JTAU) +
GAMMA * HN(JTAU))))
INTEGRATE
AINCN(M,I)
DAN(M,I) =
(AINCN - A) / GAMMA
PRINT
AINCN, DAN
4


4
N = 1
N ≤ NT
N = N + 1
Yes
No
SQG(N) =
GRADP * GRADP
SUM = 0.
N = 1
N ≤ NT - 2
N = N + 1
Yes
No
SUM = SUM +
SQG(N+1)
PNORM = .5 * SQG(1) +
SUM + 0.5 * SQG(NT)
I = 1
I ≤ NT
I = I + 1
Yes
No
ADAN(M,I)
N = 1
N ≤ NT
N = N + 1
Yes
No
PL1(N)
N = 1
N ≤ NT
N = N + 1
Yes
No
PL2(N)
SUM = 0.
N = 1
N ≤ NT - 2
N = N + 1
Yes
No
SUM = SUM +
PL1(N+1) + PL2(N+1)
QNORM = .5 * (PL1(1) + PL2(1))
+ SUM +
.5 * (PL1(NT) + PL2(NT))
RETURN TO
SUB NONCON
5


5
X(M+1)
Yes
N = 1
N ≤ NT
N = N+1
No
RETURN
TO MAIN
SQ(N) =
X(M+1,N) * X(M+1,N)
Yes
N = 1
N ≤ NT
N = N+1
No
SUM = 0.
SUM = SUM +
SQ(N+1)
Yes
N = 1
N ≤ NT-2
N = N+1
No
XNORM = .5*SQ(1) +
SUM + .5*SQ(NT)
CALL
SUB CON
Yes
IS
XNORM ≥ R
?
No
PRINT
X(M+1)
6


6
ALAM = .001
CALL SUB STEPB
GRAD1(M,I)
PNRM(M)
QNRM(M)
11
AREA = .0
I = 1
I ≤ NT
I = I + 1
Yes
PING(I)
AREA =
AREA + PING(I)
No
AREA
IS
AREA < R
?
Yes
ALAM =
ALAM + .001
IS
ALAM ≤ 1
?
Yes
GO TO 11
No
RETURN
END
No
I = 1
I ≤ NT
I = I + 1
Yes
X(M+1)
No
PRINT
X(M+1)
IS
M ≤ MT
?
Yes
GO TO 10
No
END

BIBLIOGRAPHY

VI-1. R. Aris and N.R. Amundson, Chem. Engr. Sci., Vol. 7, pp. 121-155, 1958.

VI-2. D.R. Coughanowr and L.B. Koppel, " Process Systems Analysis and Control ", Mc Graw-Hill Book Co., 1965.

VI-3. L. Lapidus, " Digital Computation for Chemical Engineers ", Mc Graw-Hill Book Co., 1962.

CHAPTER VII

CONCLUSIONS

VII - 1. Summary of Results.

The thesis investigates a design algorithm for nonlinear adaptive control systems. It is based on Kulikowski's concept of adaptive optimal control and utilizes the tools of functional analysis. The plant is described by a nonlinear vector operator equation of the integral or differential type. It is assumed that some a priori knowledge of the plant structure is available or can be obtained during an initial learning period. The unknown parameters are assumed to vary slowly compared with the total time interval required for optimization. A performance functional is specified indicative of the overall quality of plant performance and a bound is placed on the input control signal. The input and output functions are taken to be elements of an L^p-space. The problem is formulated as a conditional minimization problem. An extremum of the performance index is sought while the input function remains within or on the constraint boundary. The design objective is the construction of the input sequence whose limiting element minimizes the quality functional. The minimization problem is solved in a straightforward manner using calculus of variations techniques. The attractive and novel feature of the approach lies with the identification problem. Steady-state plant measurements and an approximate evaluation of first-order plant differentials suffice for the identification of the adjoint differentials in terms of which minimizing nonlinear operator equations are expressed. Thus, assuming asymptotically stable solutions, with a periodic input the plant output is forced to reach a steady-state and proper identification of plant differentials and their adjoints is performed. This identification step is followed by a step of optimization during which the next element of the input sequence is constructed iteratively. Identification and optimization alternate and the complexity of the overall scheme as well as the time required for each step are directly related to the plant type and its structural make-up.

Chapter III investigates the conditional optimization problem, its solution and convergence properties of the adjoint differential, and techniques for the iterative construction of the input sequence. It is shown that a hybrid scheme is needed for the implementation of the adaptive algorithm.

In Chapter IV we investigate the differential equation formulation and in Chapter V the discrete version of the problem is presented. A number of examples are taken up in Chapter VI. With the emphasis on the mathematical formulation of the adaptive algorithm, the examples are simple and intended to illustrate basic concepts only. Unimodality restrictions on the search techniques and both mathematical and computational difficulties associated with their digital implementation limit the range and scope of the examples. The computer studies of Chapter VI serve only as an initial stimulation towards the realization and eventual solution of the problems encountered in the actual implementation of the adaptive scheme.

VII - 2. Topics for Future Investigation.

During the course of this research work a number of related problems, which have not as yet found satisfactory solutions, came into focus. Topics for future research in the area of adaptive control systems include:

1. Investigation of the cases when a) the plant operator is a random operator and b) the observation of the output is disturbed by noise. The mean value of plant differentials may be obtained by a brute force method namely by averaging the results of many observations performed for the same input in consecutive time intervals. This

approach requires some additional assumptions concerning ergodicity of the random processes. Thus the basic technique introduced for the deterministic case can be utilized with some modifications but the total optimization time is greatly increased. Stochastic approximation techniques seem very promising. The basic Kiefer-Wolfowitz procedure as extended to the multidimensional case is naturally suitable to the stochastic adaptive problem. The algorithm seeks an extremum of a regression function which, in most cases previously investigated, is memoryless and unimodal. The dynamic mature of the adaptive problem and the multimodal surfaces considered necessitate some modifications of the basic algorithm.

2. Investigation of the case when the plant is described by a set of partial differential equations. The distributed parameter problem encompasses a great number of practical situations particularly in the process control industry. The adaptive solution to this problem will enhance the range of applicability of the algorithm and at the same time provide some basis for the nearly "optimal" control of more realistic processes.

3. Investigation of the "cost of control" as part of the overall index of performance along lines suggested by Feldbaum's work. And

4. Study of both deterministic and stochastic successive approximation algorithms with emphasis on techniques

that accelerate the rate of convergence of the search procedure. Digital and hybrid implementation of actual adaptive processes utilizing the theoretical results is an important segment of the topics for future research.

APPENDIX I

IDENTIFICATION OF PLANT MODEL

AI - a. Need for a Model.

In the formulation and solution of the conditional-optimization problem it is sometimes found that the iteratively constructed input lies on the boundary of the constraint. If $x_n(t)$, the n^{th} step of the input sequence, is on the constraint boundary, necessary for the construction of $x_{n+1}(t)$ is the identification of the differential of $x_n(t)$; consequently an input signal $x_n(t) + \gamma h(t)$, $t \in (0,T)$, is to be applied on-line to the plant. The time function $h(t)$ is either fixed or belongs to a set of orthonormal time functions depending on the method used for the identification of the adjoint differential operator. In practical applications the input constraint may be of either the "soft" or the "hard" type. With a "soft" constraint a tolerance band allows for an increase of the magnitude of the input signal and operation within this band may be possible by keeping γ small or taking the direction of the differential towards the constrained region rather than away from it. A "hard" constraint with $h(t)$ fixed, necessitates in most cases the use of a plant model and off-line measurement of plant differentials. Switching occurs from on-line to off-line operation whenever the input reaches the boundary of the constraint. The next member of the input sequence is evaluated using the model and this input is applied to the plant. A simple model representing the actual plant dynamics is therefore sought. Of importance is the time required for the estimation of the model parameters. What follows is a discussion of model identification schemes characterized by simplicity of representation and speed of convergence to the true plant dynamics.

AI - b. Some General Remarks.

A great many identification schemes have been proposed in the literature. These differ in three fundamental respects: the type of model used, the numerical methods used to obtain the model parameters from the measured process data, and the criterion used to determine when the model behaves like the process itself.

Two plant model types are used in this thesis. The first mathematical model used to describe the process dynamic behavior is the differential equation. According to the second the input-output relationship is given in terms of an integral operator equation. The numerical methods used to obtain the model parameters depend on the choice of model type. The criterion most commomly used to judge a model is minimum squared error between actual and model outputs when both are subjected to the same input. Other criteria used in the literature are zero error at discrete times, the Impulse Response Area Rule and the best Markov estimate. The mathematical models and the numerical methods used to obtain the model parameters are discussed below:

AI - c. Differential Equation Model.

It is assumed that the plant is described by:

$$\dot{\underline{y}}(t) = \underline{f}(\underline{x}, \underline{y}, \underline{c}) \qquad \text{(AI-1)}$$

where $\underline{y}$, $\underline{x}$, $\underline{c}$, and $\underline{f}$ are as defined in Chapter IV, section 1.

The unknown parameter vector $\underline{c}$ is to be identified. The approach to be followed for the estimation of $\underline{c}$ is the learning-model method.

Two techniques will be pursued in detail both of which have been introduced by Bellman and Kalaba [1, 2]:

a) Differential Approximation. Rewriting Equation (AI-1) as

$$\dot{\underline{y}}(t) - \underline{f}(\underline{x}, \underline{y}, \underline{c}) = 0 \quad \text{for } t \in (0,T) \qquad \text{(AI-2)}$$

it is seen that the true value of the plant parameters is the solution of:

$$\min_{\underline{c}} \int_0^T \left\{ \dot{\underline{y}} - \underline{f}(\underline{x}, \underline{y}, \underline{c}), \dot{\underline{y}} - \underline{f}(\underline{x}, \underline{y}, \underline{c}) \right\} dt \qquad \text{(AI-3)}$$

Taking partial derivatives of the integral with respect to $\underline{c}$ and setting them equal to zero yields

$$\int_0^T f_c' \dot{\underline{y}} \, dt = \int_0^T f_c' \, \underline{f} \, dt \qquad \text{(AI-4)}$$

where f_c' is a square matrix of the partial derivatives of $\underline{f}$ with respect to the vector $\underline{c}$.

The vector $\underline{c}$ is therefore evaluated by direct solution of the vector equation (AI-4).

The differential approximation method is efficient for online applications; however it is necessary to measure all the state variables $y_i(t)$ of the system and in addition some differentiation operations have to be performed on some of these variables.

b) Quasilinearization. This approach offers the distinct advantage of determining automatically together with the parameter vector $\underline{c}$ the inaccessible states of the system if any.

It is assumed that:

a) The input vector $\underline{x}(t)$ is known in an interval

$0 \leq t \leq T$. If $\underline{z}(t)$ is a subset of $\underline{y}(t)$, where $\underline{z}(t)$ is physically available for measurement, then:

b) Observations of $\underline{z}(t)$ at certain instants of time are known:

$$\underline{z}(t_i) = \alpha_i \tag{AI-5}$$

$$i = 1, 2, \ldots \ell \qquad \text{with } t_\ell \leq T$$

ℓ is chosen so that equation (AI-5) yields (m + k) conditions (m is the dimensionality of the output vector $\underline{y}(t)$ and k is the dimensionality of $\underline{c}$).

Treating the parameter vector $\underline{c}$ as a function of time ($\underline{c}$ is really a constant in the interval $0 \leq t \leq T$) and adjoining to the plant equation (AI-1) the equation:

$$\dot{\underline{c}} = 0 \tag{AI-6}$$

the problem has been reduced to the solution of the (m + k) first order equations (AI-1) and (AI-6) subject to the (m + k) boundary conditions represented by equation (AI-5).

The method of quasilinearization is used for the numerical solution of this multipoint boundary value problem: A series of approximations to the solution $\underline{y}(t)$ and

$\underline{c}$ is generated through the linearized recurrence relations:

$$\dot{\underline{y}}_{n+1} = \underline{f}(\underline{y}_n, \underline{x}, \underline{c}_n) + f_{1,n} \cdot [\underline{y}_{n+1} - \underline{y}_n] + f_{2,n} \cdot [\underline{c}_{n+1} - \underline{c}_n]$$

$$\dot{\underline{c}}_{n+1} = 0 \tag{AI-7}$$

with boundary conditions:

$$\underline{z}_{n+1}(t_1) = \alpha_1 \tag{AI-8}$$

In Equation (AI-7):

$$f_{1,n} = \left. \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{bmatrix} \right|_{\substack{\underline{y} = \underline{y}_n \\ \underline{c} = \underline{c}_n}} \tag{AI-9}$$

and:

$$f_{2,n} = \left[\begin{matrix} \frac{\partial f_1}{\partial c_1} & \cdots & \frac{\partial f_1}{\partial c_k} \\ \vdots & & \vdots \\ \frac{f_m}{c_1} & \cdots & \frac{f_m}{c_k} \end{matrix}\right]\Bigg|_{\substack{\underline{y} = \underline{y}_n \\ \underline{c} = \underline{c}_n}} \qquad \text{(AI-10)}$$

Thus the solution of equations (AI-7) and (AI-8) yields the value of the parameter vector $\underline{c}$ and the values of the inaccessible states. Computational procedures for the solution of these equations are included in References [1] and [2].

The sequence of approximations converges monotonically under certain conditions on $\underline{f}$, $f_{1,n}$ and $f_{2,n}$ and their first and second differentials, i.e. if $\underline{y}^o$ is the desired solution then:

$$\underline{y}_o < \underline{y}_1 < \cdots < \underline{y}^o$$

Also the convergence is quadratic, i.e.

$$\left| Y_{n+1} - Y^{o} \right| \leq K \left| Y_{n} - Y^{o} \right|^{2} \qquad k \leq 1$$

a computation advantage resulting in rapid convergence.

Explicit mathematical-relation methods as well as algorithms based on a modified Newton's procedure for the estimation of plant parameters are described in the literature [3].

AI - d. Integral Equation Model.

Here the mathematical relation for the system that relates the output to the input is an integral equation. A number of approaches have been suggested for the solution of the parameter estimation problem assuming the above process topology to be known.

Ku and Wolf [4] give recurrence relations for the Volterra Kernels. Their algorithm is rather cumbersome to apply in practical situations and the impulse response of the linear part of the process is assumed to be known.

Hsieh [5] used a computational algorithm based on the steepest descent method, which was previously employed by Balakrishnan for solving a class of filtering and control problems, in order to estimate the various order weighting function matrices through the observation of input and output data of the nonlinear process over a finite time interval. A least square criterion was used and an updating scheme was devised based on steepest descent as more data were available.

Eykoff [6] suggested the following procedure: Consider the scheme in Fig. (AI-2). Process P is nonlinear and has dynamic properties. So model M has to be nonlinear too. Only the time invariant case is considered. The model output z(t) is given by:

$$z(t) = \int_{-\infty}^{t} x(\tau)h_1(t-\tau)d\tau + \int_{-\infty}^{t}\int_{-\infty}^{t} x(\tau_1)x(\tau_2)h_2(t-\tau_1, t-\tau_2)d\tau_1 d\tau_2 \qquad \text{(AI-11)}$$

The series (AI-11) can be approximated by finite sums if the process is stable, i.e. has finite memory. Then:

$$z(t) \approx \sum_i \alpha_i x(t-\tau_i) + \sum_i \sum_j \alpha_{ij} x(t-\tau_i)x(t-\tau_j) + \sum_i \sum_j \sum_k \cdots \qquad \text{(AI-12)}$$

Now consider the diagrammatical representation of Fig. (AI-3), where P is again the process, N is a nonlinear filter the operation of which is given by equation (AI-11) and T_d is a time delay operator. Minimization of the integral squared error E, where:

$$E = \int_0^T e^2 \, dt \qquad \text{(AI-13)}$$

and $e = y - z$ results in the following condition for everyone of the coefficients α :

$$\frac{\partial E}{\partial \alpha} \rightarrow 0 \tag{AI-14}$$

If $u_i \triangleq x(t - \tau_i)$

and defining

$$\xi_i = \frac{1}{2} \frac{\partial E}{\partial \alpha_i} \equiv \langle e, u_i \rangle$$

$$\xi_{ij} = \frac{1}{2} \frac{\partial E}{\partial \alpha_{ij}} \equiv \langle e, u_i u_j \rangle$$

etc.

as the components of the error vector, the mechanism involves adjustment of the coefficients α so that the error vector becomes equal to zero.

A mechanization of this model is shown in Fig. (AI-4) where the symbol $\langle\ \rangle$ indicates the inner product of the two input quatities. In general a time varying process will be assumed to be piecewise time invariant so that a model of the above configuration be possible. For greater or less accurate representation of the actual process dynamics a more complicated or correspondingly simpler model structure is available. Details of model implementation schemes can be found in the references.

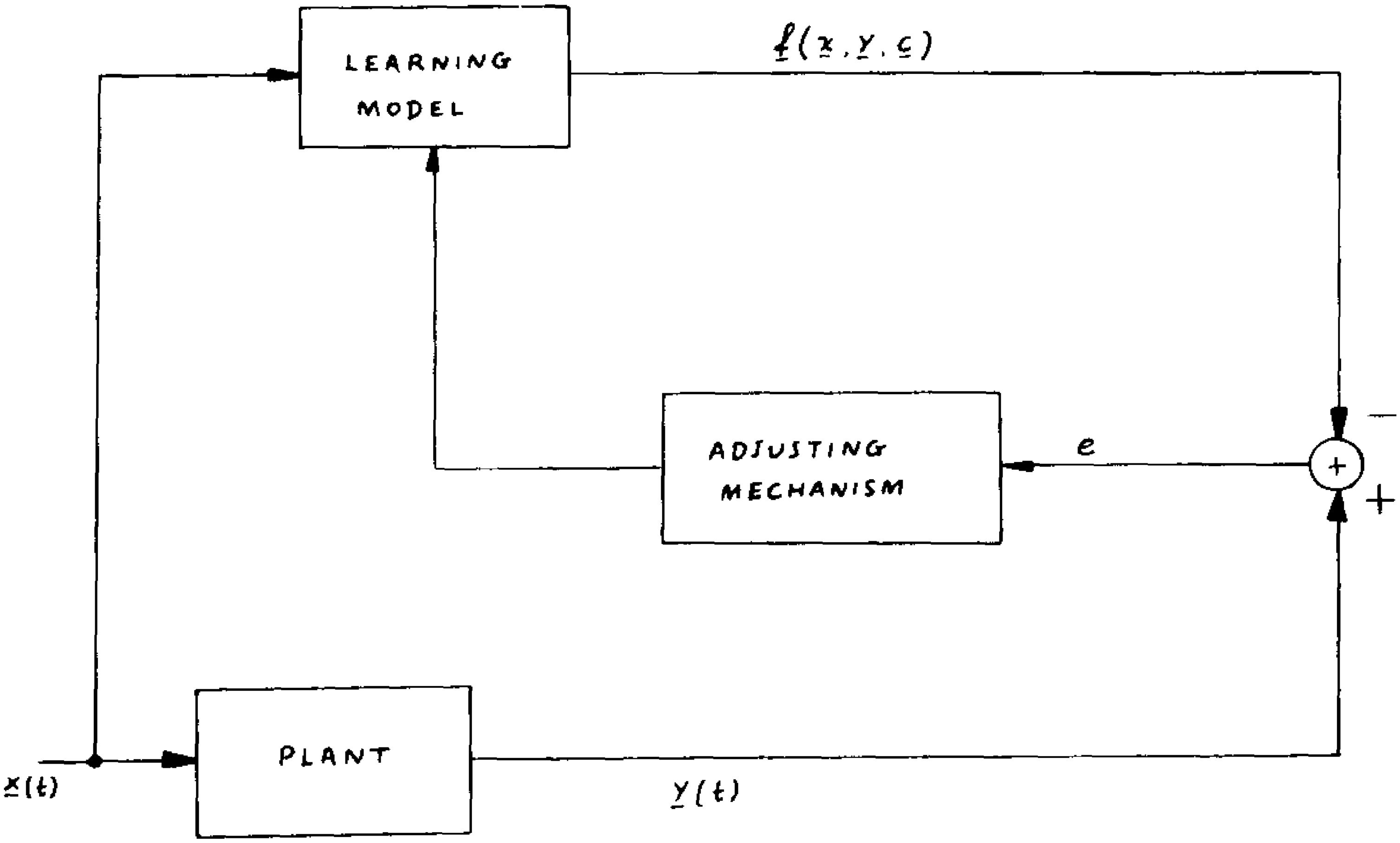


FIGURE AI - 1

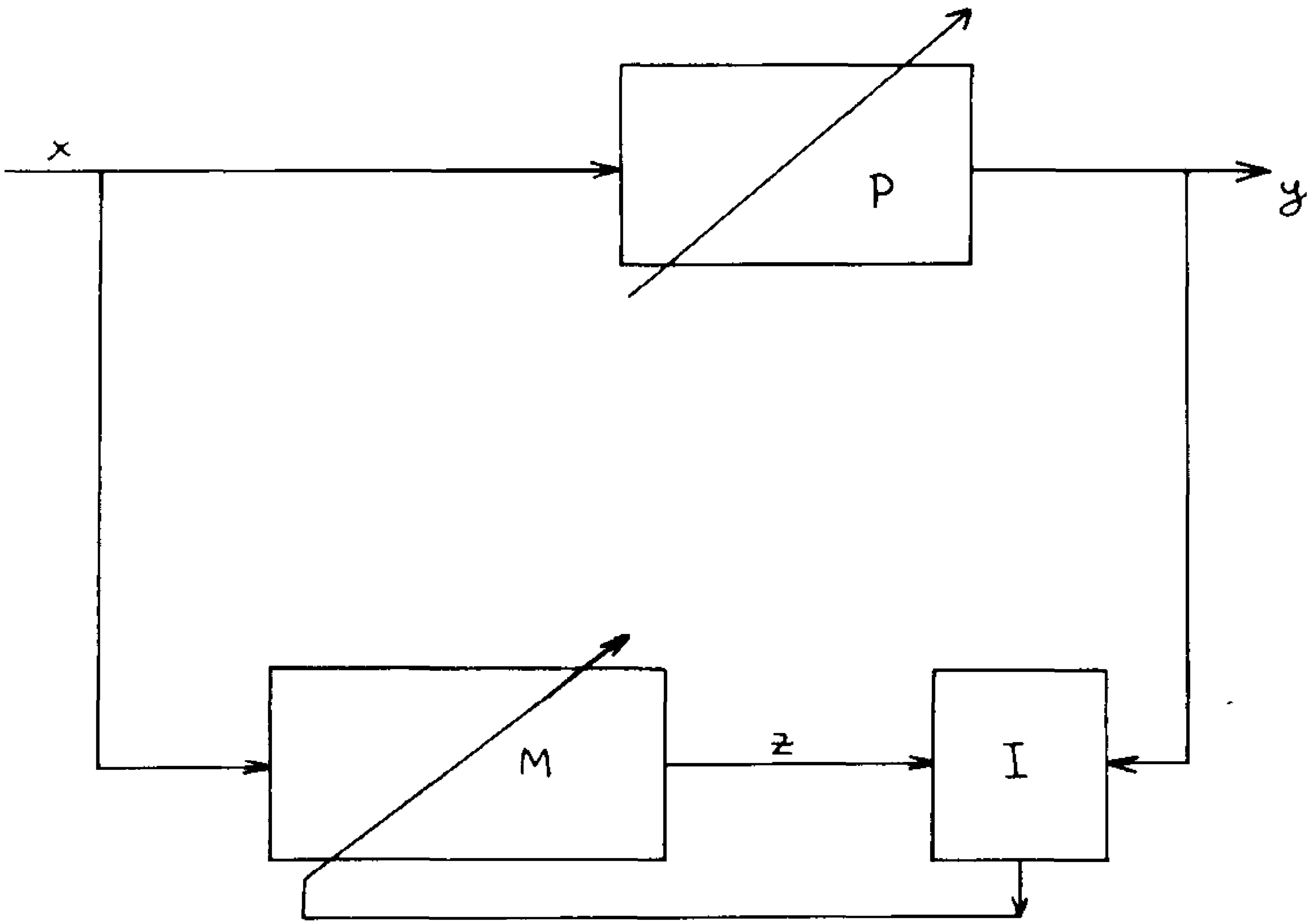


FIGURE AI-2

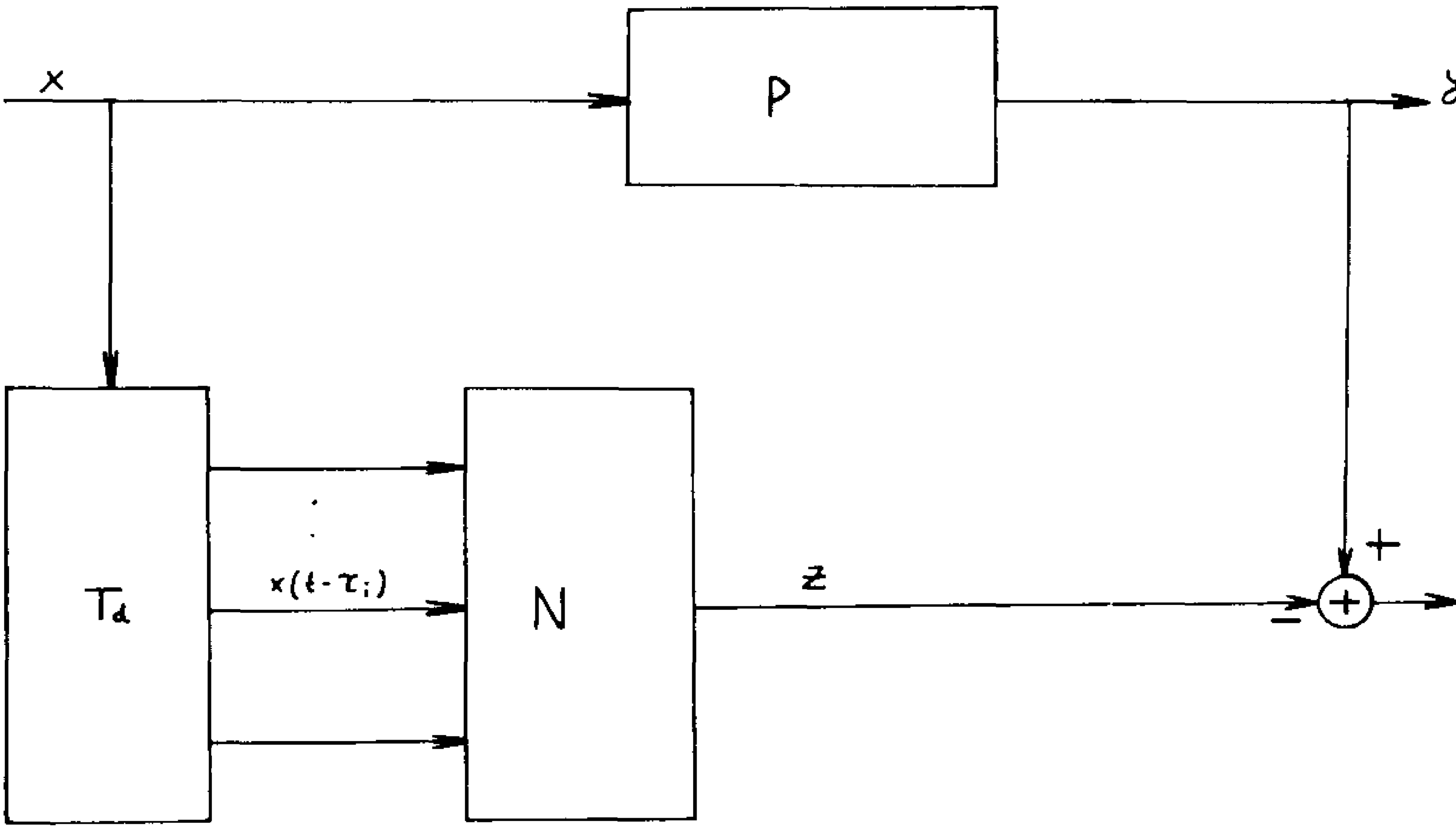
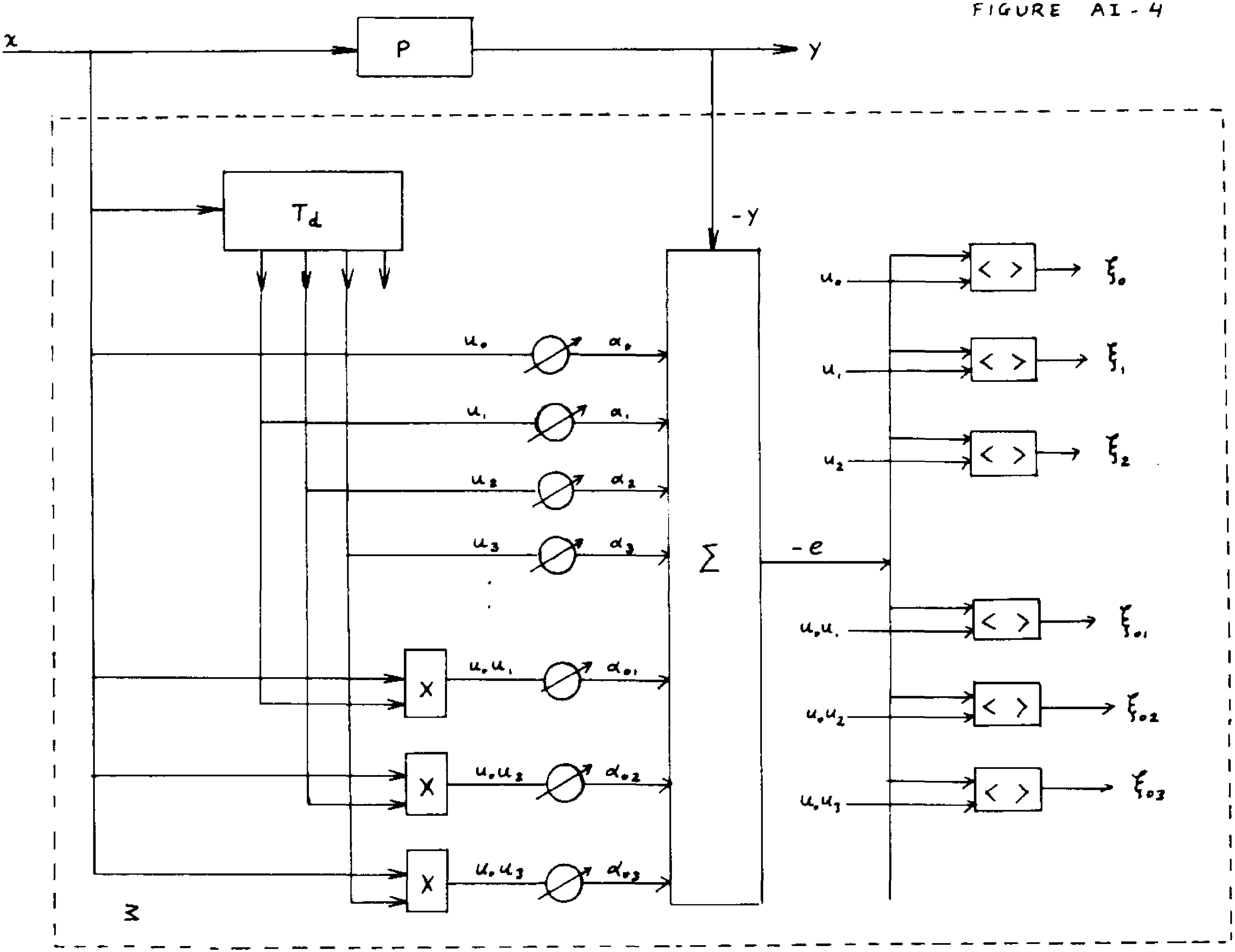


FIGURE AI - 3

FIGURE AI-4


x
P
y
-y
T_d
u_0
u_1
u_2
u_3
α_0
α_1
α_2
α_3
u_0u_1
u_0u_2
u_0u_3
α_{01}
α_{02}
α_{03}
X
Σ
-e
ξ_0
ξ_1
ξ_2
ξ_{01}
ξ_{02}
ξ_{03}
M

BIBLIOGRAPHY

1. R.E. Bellman and R.E. Kalaba,"Quasilinearization and Nonlinear Boundary Value Problems." New York/ American Elsevier Publishing Co. , 1965.

2. R.E. Bellman, R.E. Kalaba and R. Sridhar, "Adaptive Control Via Quasiliearization and Differential Approximation", Computing Vo. 1, Fasc. 1, 1966.

3. D.D. Donalson and F.H. Kishi "Review of Adaptive Control System Theories and Techniques" in Modern Control System Theory edited by C.T. Leondes, pp. 271-281.

4. Y.H. Ku and A.A. Wolf, "Volterra-Weiner Functionals for the Analysis of Nonlinear Systems" , Journal of the Franklin Institute , Vol. 281, No. 1, Jan. 1966, pp. 9-26.

5. H.C. Hsieh, "An On-Line Identification Scheme for Multivariable Nonlinear Systems", Proceedings of Conference of Computing Methods in Optimization Problems, Univ. of Californa, Los Angeles, edited by Balakrishnan and Neustadt, pp. 193-210, 1964.

6. P. Eykoff, "Process Parameter Estimation", in Progress in Control Engineering, Vol. 2 Academic Press, 1964.

APPENDIX II

This appendix contains a listing of the computer programs for the case studies examined in this thesis and referred to in Chapter VI.

```
C     THIS PROGRAM EVALUATES ITERATIVELY A SEQUENCE OF CONTROL FUNCTIONS
C     FOR A REGULATOR ADAPTIVE CONTROL PROBLEM.
C
C     THE PLANT IS THE CASCADE OF A NONLINEAR AND A LINEAR PART.
C
C     THE PLANT OPERATOR IS OF THE INTEGRAL TYPE.
C
C     THE PERFORMANCE INDEX IS THE INTEGRAL SQUARE OF THE DIFFERENCE
C     BETWEEN DESIRED AND ACTUAL PLANT OUTPUTS.
C
C     THE ENERGY OF THE INPUT FUNCTION IS CONSTRAINED TO BE LESS OR
C     EQUAL TO A CONSTANT VALUE.
C
C     THE REVERSE OPERATOR METHOD IS UTILIZED FOR THE SOLUTION OF THE
C     IDENTIFICATION PROBLEM.
C
C     THE OPTIMAL INPUT SEQUENCE IS OBTAINED VIA ALTMAN'S ITERATION
C     TECHNIQUE.
C
C
C     VARIABLES USED IN THIS PROGRAM:
C
C     X    -CONTROL VARIABLE
C     A    -PLANT OUTPUT
C     AI, AINC, AINCN    -PLANT OUTPUT WITH FIXED INPUT VARIATION
C     PLNT1    -INTEGRAND OF PLANT OUTPUT EXPRESSION
C     PLNT2, PLNT3, PLNT4    -INTEGRAND OF PLANT OUTPUT EXPRESSION
C                             WITH A FIXED INPUT VARIATION
C     H, HN, HNI    -FIXED INPUT VARIATIONS
C     G    -THE DIFFERENCE BETWEEN DESIRED AND ACTUAL PLANT OUTPUTS
C     DA, DAN, DANI    -PLANT OUTPUT DIFFERENTIALS
C     ADA, ADAN, ADANI    -THE ADJOINTS OF THE PLANT DIFFERENTIALS
C     XNORM    -THE NORM OF THE INPUT VARIABLE
C     DNORM    -THE NORM OF THE PERFORMANCE CRITERION
C     GRADP    -THE GRADIENT OF THE PERFORMANCE CRITERION
C     GRADI    -THE GRADIENT OF THE MODIFIED PERFORMANCE CRITERION
C
C     THE REST OF THE VARIABLES USED ARE COMBINATIONS OF THE ABOVE
C     GENERATED SOLELY FOR CONVENIENCE.
C
C
C     CONSTANTS USED IN THIS PROGRAM:
C
C     MT    -TOTAL NUMBER OF ITERATION STEPS
C     NT    -NUMBER OF SAMPLING POINTS IN THE OPTIMIZATION INTERVAL
C     ALPHA, BETA, C    -CONSTANTS ASSOCIATED WITH THE PLANT STRUCTURE
C     YD    -DESIRED OUTPUT LEVEL
C     GAMMA    -PARAMETER USED IN THE EVALUATION OF THE PLANT
C               DIFFERENTIAL
C     R    -VALUE OF INPUT CONSTRAINT
C
C     ****************************************************************
C     *
C
C     THIS IS THE MAIN PROGRAM
C
C     IT SETS THE VALUES OF THE CONSTANT PARAMETERS, CALLS THE
C     SUBROUTINES AND CHECKS FOR THE INPUT CONSTRAINT.
C
C
```

```
      DIMENSION X(101,11),A(101,11),AINC(101,11),DA(101,11),ADA(101,11),
     1XNORM(101),SQ(11),H(11),AO(11),YD(11)
      COMMON X,A,AINC,DA,ADA,XNORM,YD,H,M,N,MT,NT,ALPHA,BETA,GAMMA,C,R,
     1AO,ZX,ZYD,ZAO,SW
C
C     THE FOLLOWING STATEMENTS SET THE INITIAL CONDITIONS AND THE
C     DESIRED OUTPUT LEVEL.
C
      READ 1, MT,NT
      READ 2, ALPHA,BETA,GAMMA,C,R
      READ 3, ZX,ZYD,ZAO
      DO 4 L=1,NT
      X(1,L)=ZX
      YD(L)=ZYD
    4 AO(L)=ZAO
    1 FORMAT (2I10)
    2 FORMAT (5F10.4)
    3 FORMAT (3F10.4)
      SW=0.0
      DO 1000 M=1,MT
      CALL PLANT
      CALL ADJUNT
      CALL NONCON
C
C     THIS PART OF THE PROGRAM EVALUATES THE NORM OF THE INPUT FUNCTION
C     AND TESTS THE CONSTRAINT CONDITION.
C
      IF (SW-1.0) 500,2000,2000
  500 DO 100 N=1,NT
  100 SQ(N)=X(M+1,N)*X(M+1,N)
      SUM=0.0
      NT2=NT-2
      DO 200 N=1,NT2
  200 SUM=SUM+SQ(N+1)
      XNORM(M+1)=0.5*SQ(1)+SUM+0.5*SQ(NT)
      IF (XNORM(M+1)-R) 800,900,900
  900 CALL CON
      IF (SW-1.0) 800,2000,2000
  800 MP=M+1
C
C     THE FOLLOWING STATEMENT PRINTS THE VALUES OF THE CONTROL FUNCTION
C     OBTAINED FROM THE PREVIOUS ITERATION STEP.
C
      PRINT 7, (MP,N,X(MP,N), N=1,NT)
    7 FORMAT (1H ,/(15X,2HX(,I3,1H,,I2,3H) =,E13.4,/))
 1000 CONTINUE
 2000 STOP
      END
C
C     ****************************************************************
C
      SUBROUTINE PLANT
C
C
C     THIS SUBROUTINE EVALUATES THE OUTPUT ARRAY A(M,N) AND THE FIRST
C     PLANT DIFFERENTIAL DA(M,N).
C
      DIMENSION X(101,11),A(101,11),AINC(101,11),DA(101,11),ADA(101,11),
     1XNORM(101),SQ(11),H(11),AO(11),YD(11)
     2,PENT1(11),PENT2(11),G(101,11),REVG(101,11)
```

```
      COMMON X,A,AINC,DA,ADA,XNORM,YD,H,M,N,MI,NI,ALPHA,BETA,GAMMA,C,K,
     1AO,ZX,ZYD,ZAO,SW
    1 FORMAT (1H ,4H  M=I3,2X,2HI=I2,2X,7HA(M,I)=F12.3,2X,5HH(I)=F12.3,
     12X,10HAINC(M,I)=F12.3,2X,8HDA(M,I)=F12.3)
    3 FORMAT (1H ,16X,7HA(M,1)=F12.3)
      A(M,1)=ZAO
      PRINT 3, A(M,1)
C
C     A TRAPEZOIDAL RULE IS USED TO EVALUATE THE PLANT INTEGRALS.
C
      DO 100 I=2,NI
      DO 200 JTAU=1,I
      IMTAU=I-JTAU
  200 PLNT1(JTAU)=C*EXP(-BETA*(IMTAU))*(1.0-EXP(-ALPHA*X(M,JTAU)))
      IF (I-2) 210,210,220
  210 A(M,I)=0.5*PLNT1(1)+0.5*PLNT1(2)
      GO TO 110
  220 SUM1=0.0
      IM2=I-2
      DO 250 J=1,IM2
  250 SUM1=SUM1+PLNT1(J+1)
      A(M,I)=0.5*PLNT1(1)+SUM1+0.5*PLNT1(I)
  110 A(M,I)=ZAO+A(M,I)
  100 CONTINUE
C
C     THE FOLLOWING STATEMENTS EVALUATE THE FIXED INPUT VARIATION H(I).
C
      DO 205 I=1,NI
  205 G(M,I)=YD(I)-A(M,I)
      DO 307 I=1,NI
      REVG(M,I)=G(M,NI+1-I)
  307 H(I)=EXP(ALPHA*X(M,I))*REVG(M,I)
C
C     AN APPROXIMATE ESTIMATE FOR THE PLANT DIFFERENTIAL DA(M,N) IS
C     OBTAINED BELOW.
C
      AINC(M,1)=ZAO
      DA(M,1)=0.0
      DO 300 I=2,NI
      DO 400 JTAU=1,I
  400 PLNT2(JTAU)=C*EXP(-BETA*(I-JTAU))*(1.0-EXP(-ALPHA*(X(M,JTAU)+
     1GAMMA*H(JTAU))))
      IF (I-2) 215,215,225
  215 AINC(M,I)=0.5*PLNT2(1)+0.5*PLNT2(2)
      GO TO 115
  225 SUM2=0.0
      IM2=I-2
      DO 255 J=1,IM2
  255 SUM2=SUM2+PLNT2(J+1)
      AINC(M,I)=0.5*PLNT2(1)+SUM2+0.5*PLNT2(I)
  115 AINC(M,I)=ZAO+AINC(M,I)
      DA(M,I)=(AINC(M,I)-A(M,I))/GAMMA
  300 PRINT 1, M,I,A(M,I),H(I),AINC(M,I),DA(M,I)
      RETURN
      END
C
C     ************************************************************
      SUBROUTINE ADJUNT
C
C
```

```
C     SUBROUTINE ADJOINT EVALUATES THE ADJOINT PLANT DIFFERENTIAL
C     USING THE REVERSE OPERATOR METHOD.
C
C
      DIMENSION X(101,11),A(101,11),AINC(101,11),DA(101,11),ADA(101,11),
     1XNORM(101),SQ(11),H(11),AO(11),YD(11)
     2,REVDA(101,11)
      COMMON X,A,AINC,DA,ADA,XNORM,YD,H,M,N,MT,NT,ALPHA,BETA,GAMMA,C,K,
     1AO,ZX,ZYD,ZAO,SW
    1 FORMAT (1H ,15X,4HADA(,I3,1H,,I2,4H) = ,E13.4)
      DO 100 I=1,NT
      REVDA(M,I)=DA(M,NT+1-I)
      ADA(M,I)=EXP(-ALPHA*X(M,I))*REVDA(M,I)
  100 PRINT 1, M,I,ADA(M,I)
      RETURN
      END
C
C     ******************************************************************
      SUBROUTINE NONCON
C
C
C
C     THIS SUBROUTINE EVALUATES THE NORM OF THE DIFFERENCE FUNCTION
C     YD - A. IF THE VALUE OF THE NORM LIES WITHIN AN ACCEPTABLE
C     TOLERANCE BAND THE PROGRAM IS TERMINATED. OTHERWISE, THE NEXT
C     ELEMENT OF THE CONTROL SEQUENCE IS ESTIMATED VIA ALTMAN'S TECHNIQUE
C     FOR THE UNCONSTRAINED PROBLEM.
C
      DIMENSION X(101,11),A(101,11),AINC(101,11),DA(101,11),ADA(101,11),
     1XNORM(101),SQ(11),H(11),AO(11),YD(11)
     2,SQD(11),DNORM(101)
     3,GRADP(101,11),AINCN(101,11),DAN(101,11),ADAN(101,11),HN(11),
     4PENES(11),SQG(11),PNORM(101),GNORM(101),PL1(11),PL2(11)
      COMMON X,A,AINC,DA,ADA,XNORM,YD,H,M,N,MT,NT,ALPHA,BETA,GAMMA,C,K,
     1AO,ZX,ZYO,ZAO,SW
     2,GRADP,AINCN,DAN,ADAN,HN,PENES,SQG,PNORM,GNORM,PL1,PL2
      SUMSQD=0.0
      DO 100 I=1,NT
      SQD(I)=(YD(I)-A(M,I))*(YD(I)-A(M,I))
  100 SUMSQD=SUMSQD+SQD(I)
      DNORM(M)=SQRT(SUMSQD)
      IF (DNORM(M)-0.01) 400,400,500
  400 SW=1.0
      RETURN
  500 CALL STEPA
C
C
C     SUBROUTINE STEPA CALCULATES THE VALUES OF THE PARAMETERS
C     UTILIZED IN ALTMAN'S ITERATION METHOD.
C
      DO 200 I=1,NT
  200 X(M+1,I)=X(M,I)-(PNORM(M)*GRADP(M,I))/2.*GNORM(M)
      RETURN
      END
C
C     ******************************************************************
      SUBROUTINE STEPA
C
C
C
C     THIS SUBROUTINE CALCULATES THE GRADIENT OF THE PERFORMANCE
C     FUNCTIONAL AND APPROPRIATE NORMS FOR THE IMPLEMENTATION OF
C     ALTMAN'S PROCEDURE.
```

```
C
      DIMENSION X(101,11),A(101,11),AINC(101,11),DA(101,11),ADA(101,11),
     1XNORM(101),SQ(11),H(11),AO(11),YD(11)
     2,SQD(11),ENORM(101)
     3,GRADP(101,11),AINCN(101,11),DAN(101,11),ADAN(101,11),HN(11),
     4PLNT3(11),SQG(11),PNORM(101),QNORM(101),PL1(11),PL2(11)
      COMMON X,A,AINC,DA,ADA,XNORM,YD,B,M,N,MT,NT,ALPHA,BETA,GAMMA,C,R,
     1AO,ZX,ZYD,ZAO,SN
     2,GRADP,AINCN,DAN,ADAN,HN,PLNT3,SQG,PNORM,QNORM,PL1,PL2
C
C     THE FOLLOWING STATEMENT EVALUATES THE GRADIENT OF THE PERFORMANCE
C     FUNCTIONAL.
C
      DO 501 I=1,NT
  501 GRADP(M,I)=-2.*ADA(M,I)
C
C     THE FIXED INPUT VARIATION IS SET EQUAL TO GRADP.
C
      DO 10 I=1,NT
   10 HN(I)=GRADP(M,I)
C
C     THE PLANT DIFFERENTIAL IS EVALUATED BELOW.
C
      AINCN(M,1)=0.
      DAN(M,1)=0.
      DO 30 I=2,NT
      DO 40 JTAU=1,I
   40 PLNT3(JTAU)=C*EXP(-BETA*(I-JTAU))*(1.-EXP(-ALPHA*(X(M,JTAU)+
     1GAMMA*HN(JTAU))))
      IF(I-2) 21,21,22
   21 AINCN(M,I)=0.5*PLNT3(1)+0.5*PLNT3(2)
      GO TO 11
   22 SUM2=0.
      IM2=I-2
      DO 25 J=1,IM2
   25 SUM2=SUM2+PLNT3(J+1)
      AINCN(M,I)=0.5*PLNT3(1)+SUM2+0.5*PLNT3(I)
   11 DAN(M,I)=(AINCN(M,I)-A(M,I))/GAMMA
   30 PRINT 1, M,I,AINCN(M,I),DAN(M,I)
    1 FORMAT (1H ,4H  M=I3,2X,2HI=I2,2X,11HAINCN(M,I)=F12.3,2X,
     19HDAN(M,I)=F12.3)
C
C     THIS PART OF THE PROGRAM EVALUATES VIA THE TRAPEZOIDAL RULE THE
C     NORM OF THE GRADIENT OF THE PERFORMANCE CRITERION.
C
      DO 50 N=1,NT
   50 SQG(N)=GRADP(M,N)*GRADP(M,N)
      SUM=0.
      NT2=NT-2
      DO 60 N=1,NT2
   60 SUM=SUM+SQG(N+1)
      PNORM(M)=0.5*SQG(1)+SUM+0.5*SQG(NT)
C
C     THE ADJOINT DIFFERENTIAL IS EVALUATED USING THE REVERSE OPERATOR
C     METHOD.
C
      DO 65 I=1,NT
   65 ADAN(M,I)=EXP(-ALPHA*X(M,I))*DAN(M,NT+1-I)
C
C     THIS PART EVALUATES THE NORM WHICH APPEARS IN ALTMAN'S EXPRESSION.
```

```
      DO 70 N=1,NT
   70 PL1(N)=2.*DA(M,N)*AN(I,N)
      DO 80 N=1,NT
   80 PL2(N)=2.*ALPHA*GRADP(M,N)*GRADP(M,N)*ADAN(M,N)
      SUM=0.
      NT2=NT-2
      DO 85 N=1,NT2
   85 SUM=SUM+PL1(N+1)+PL2(N+1)
      QNORM(M)=0.5*(PL1(1)+PL2(1))+SUM+0.5*(PL1(NT)+PL2(NT))
      RETURN
      END

      ******************************************************************
      SUBROUTINE CON

      THIS SUBROUTINE ITERATIVELY GENERATES THE CONTROL SEQUENCE VIA
      ALTMAN'S TECHNIQUE FOR THE CONSTRAINED PROBLEM.

      DIMENSION X(101,11),A(101,11),AINC(101,11),DA(101,11),ADA(101,11),
     1XEGRE(101),SP(11),H(11),AG(11),YD(11)
     2,PING(11)
     3,GRADI(101,11),PNI(11),AI(101,11),DANI(101,11),PLN14(11),
     4SQGL(11),ERRG(101),CNRM(101),ADANI(101,11),PL3(11),PL4(11)
      COMMON X,A,AINC,DA,ADA,XNORM,YD,H,M,N,MT,NT,ALPHA,BETA,GAMMA,C,R,
     1AG,ZX,ZYD,ZAD,SP
     2,GRADI,PNI,AI,DANI,PLN14,SQGL,PNRM,CNRM,ADANI,PL3,PL4
C
C     A VALUE FOR THE WEIGHTING PARAMETER ALAM WILL BE CHOSEN BY TESTING
C     SO THAT THE CONTROL SIGNAL LIES ON THE BOUNDARY OF THE CONSTRAINT.
C
      ALAM=0.001
      CALL STEPB
C
C     SUBROUTINE STEPB EVALUATES NECESSARY NORMS AND FUNCTIONS FOR THE
C     APPLICATION OF ALTMAN'S TECHNIQUE.THE CONSTRAINED PROBLEM IS
C     CONSIDERED.
C
C
C     THE FOLLOWING PART EVALUATES THE NORM OF THE INPUT AND COMPARES IT
C     WITH THE VALUE OF THE CONSTRAINT.
C
   30 AREA=0.0
      DO 25 I=1,NT
      PING(I)=(X(M,I)-(PNRM(M)*GRADI(M,I))/2.*CNRM(M))*
     1(X(M,I)-(PNRM(M)*GRADI(M,I))/2.*CNRM(M))
   25 AREA=AREA+PING(I)
      AREA=AREA-0.5*(PING(1)+PING(NT))
      IF (AREA-R) 10,20,20
   10 ALAM=ALAM+0.001
      IF (ALAM.LE.1.0) GO TO 30
      ALAM=0.0
      PRINT 920, ALAM
  920 FORMAT (1H ,5X,'ALAM=',F6.3)
      RETURN
C
C     THIS PART GENERATES THE CONTROL SEQUENCE.
C
   20 DO 21 I=1,LT
```

```
   21 X(M+1,I)=X(M,I)-(PNEM(M)*GRADI(M,I))/2.*GNRM(M)
      PRINT 910, ALAM
  910 FORMAT (1H ,5X,'ALAM=',F6.3)
      RETURN
      END
C
C     **********************************************************************
      SUBROUTINE STEP3
C
C
C     THIS SUBROUTINE CALCULATES THE GRADIENT OF THE MODIFIED
C     PERFORMANCE CRITERION AND APPROPRIATE NORMS FOR THE IMPLEMENTATION
C     OF ALTMAN'S PROCEDURE. THE CONSTRAINED PROBLEM IS CONSIDERED.
C
      DIMENSION X(101,11),A(101,11),AINC(101,11),DA(101,11),ADA(101,11),
     1XNORM(101),SO(11),H(11),AO(11),YO(11)
     2,PINS(11)
     3,GRADI(101,11),HNI(11),AI(101,11),DANI(101,11),PLNI4(11),
     4SQGI(11),PNRM(101),GNRM(101),ADANI(101,11),PL3(11),PL4(11)
      COMMON X,A,AINC,DA,ADA,XNORM,YO,H,M,N,IT,NT,ALPHA,BETA,GAMMA,C,R,
     1AO,ZX,ZYO,ZAO,SN
     2,GRADI,HNI,AI,DANI,PLNI4,SQGI,PNRM,GNRM,ADANI,PL3,PL4
C
C     THE GRADIENT OF THE MODIFIED PERFORMANCE FUNCTIONAL IS EVALUATED
C
      DO 501 I=1,NT
  501 GRADI(M,I)=2.*ALAM*X(M,I)-2.*ADA(M,I)
C
C     THE FIXED INPUT VARIATION IS SET EQUAL TO GRADI
C
      DO 10 I=1,NT
   10 HNI(I)=GRADI(M,I)
C
C     THE PLANT DIFFERENTIAL IS EVALUATED BELOW
C
      AI(M,1)=0.
      DANI(M,1)=0.
      DO 30 I=2,NT
      DO 40 JTAU=1,I
   40 PLNI4(JTAU)=C*EXP(-BETA*(I-JTAU))*(1.-EXP(-ALPHA*(X(M,JTAU)+
     1GAMMA*HNI(JTAU))))
      IF (I-2) 21,21,22
   21 AI(M,I)=0.5*PLNI4(1)+0.5*PLNI4(2)
      GO TO 11
   22 SUM2=0.
      IM2=I-2
      DO 25 J=1,IM2
   25 SUM2=SUM2+PLNI4(J+1)
      AI(M,I)=0.5*PLNI4(1)+SUM2+0.5*PLNI4(I)
   11 DANI(M,I)=(AI(M,I)-A(M,I))/GAMMA
   30 PRINT 1, M,I,AI(M,I),DANI(M,I)
    1 FORMAT (1H ,4H  M=I3,2X,2HI=I2,2X,8HAI(M,I)=F12.3,2X,
     110HDANI(M,I)=F12.3)
C
C     THIS PART OF THE PROGRAM EVALUATES VIA THE TRAPEZOIDAL RULE THE
C     NORM OF THE GRADIENT OF THE MODIFIED PERFORMANCE CRITERION.
C
      DO 50 N=1,NT
   50 SQGI(N)=GRADI(M,N)*GRADI(M,N)
      SUM=0.
```

```
      NI2=NI-2
      DO 60 N=1,NI2
   60 SUM=SUM+SQGI(N+1)
      PERM(M)=0.5*SQGI(1)+SUM+0.5*SQGI(NI)
C
C     THE ADJOINT DIFFERENTIAL IS EVALUATED USING THE REVERSE OPERATOR
C     METHOD.
C
      DO 65 I=1,NI
   65 ADAN1(M,I)=EXP(-ALPHA*X(M,I))*DAN1(M,NI+1-I)
C
C     THIS PART EVALUATES THE NORM WHICH APPEARS IN ALTHAN'S EXPRESSION.
C
      DO 70 N=1,NI
   70 PL3(N)=2.*(GRAD1(M,N)*GRAD1(M,N)*ALAM+DAN1(M,N)*DAN1(M,N))
      DO 80 N=1,NI
   80 PL4(N)=2.*ALPHA*GRAD1(M,N)*GRAD1(M,N)*ADAN1(M,N)
      SUM=0.
      NI2=NI-2
      DO 85 N=1,NI2
   85 SUM=SUM+PL3(N+1)+PL4(N+1)
      QNORM(M)=0.5*(PL3(1)+PL4(1))+SUM+0.5*(PL3(NI)+PL4(NI))
      RETURN
      END
C
C     ****************************************************************
```

```
C      THIS PROGRAM EVALUATES ITERATIVELY A SEQUENCE OF CONTROL FUNCTIONS
C      FOR A PLANT CONSISTING OF THE PARALLEL CONNECTION OF A LINEAR
C      INTEGRATOR AND A NONLINEAR PART.
C
C      THE ACCUMULATED PLANT OUTPUT IN A FIXED TIME INTERVAL IS MAXIMIZED.
C
C      THE ENERGY OF THE CONTROL SIGNAL CONSTRAINED TO BE LESS OR
C      EQUAL TO A CONSTANT VALUE.
C
C      THE REVERSE OPERATOR METHOD IS UTILIZED FOR THE SOLUTION OF THE
C      IDENTIFICATION PROBLEM.
C
C      THE OPTIMAL INPUT SEQUENCE IS OBTAINED VIA NEWTON'S METHOD WITH
C      A FIXED STEPSIZE.
C
C      THE OPTIMIZATION INTERVAL IS DIVIDED INTO TWENTY EQUAL
C      SUBINTERVALS.
C
C      VARIABLES USED IN THIS PROGRAM:
C
C      X    -CONTROL VARIABLE
C      A    -PLANT OUTPUT
C      PINT1    -INTEGRAND OF PLANT OUTPUT EXPRESSION
C      PINT3    -INTEGRAND OF PLANT OUTPUT EXPRESSION WITH A FIXED INPUT
C      VARIATION
C      H    -FIXED INPUT VARIATION
C      G    -THE VARIATION OF THE ADJOINT DIFFERENTIAL EQUAL TO UNITY
C      DAD    -PLANT OUTPUT DIFFERENTIAL
C      ADA    -THE ADJOINT OF THE PLANT DIFFERENTIAL
C      XNORM    -THE NORM OF THE INPUT VARIABLE
C
C      CONSTANTS USED IN THIS PROGRAM:
C
C      MT    -TOTAL NUMBER OF ITERATION STEPS
C      NT    -NUMBER OF SAMPLING POINTS IN THE OPTIMIZATION INTERVAL
C      T    -OPTIMIZATION PERIOD
C      EPS    -WEIGHTING FACTOR OF THE NONLINEAR PART
C      R    -VALUE OF INPUT CONSTRAINT
C
C      THIS IS THE MAIN PROGRAM
C      IT SETS THE VALUES OF THE CONSTANT PARAMETERS, THE INITIAL
C      CONDITIONS, CALLS THE SUBROUTINES AND CHECKS FOR THE INPUT
C      CONSTRAINT.
C
       DIMENSION X(101,21),A(101,21),DAD(101,21),ADA(101,21),XNORM(101),
      1PINGD(21),H(21)
       COMMON X,A,DAD,ADA,XNORM,PINGD,H,M,N,MT,NT,T,EPS,R,SW,XO
       READ 1, MT,NT
       READ 3, T,EPS,R
       READ 5, (X(1,N), N=1,10)
       READ 9, (X(1,N), N=11,21)
     1 FORMAT (2I10)
     3 FORMAT (3F10.4)
     5 FORMAT (10F4.2)
     9 FORMAT (11F4.2)
       SW=0.0
       DO 1000 M=1,MT
       CALL PLANT
       CALL ADJOINT
```

```
      CALL NORCON
      IF (SW-1.0) 500,2000,2000
C
C     THIS PART OF THE PROGRAM EVALUATES THE NORM OF THE INPUT FUNCTION
C     AND TESTS THE CONSTRAINT CONDITION.
C
  500 DO 200 JT=1,NT
  200 PINGD(JT)=X(M+1,JT)*X(M+1,JT)
      SUM=0.0
      NTM2=NT-2
      DO 300 J=1,NTM2
  300 SUM=SUM+PINGD(J+1)*0.2
      XNORM(M+1)=0.5*PINGD(1)*0.2+SUM+0.5*PINGD(NT)*0.2
      IF (XNORM(M+1)-R) 800,900,900
  900 CALL CON
C
C     SUBROUTINE CON IS CALLED WHENEVER THE NORM OF THE CONTROL SIGNAL
C     EXCEEDS THE VALUE OF THE CONSTRAINT R.
C
  800 MP=M+1
C
C     THE FOLLOWING STATEMENT PRINTS THE VALUES OF THE CONTROL FUNCTION
C     OBTAINED FROM THE PREVIOUS ITERATION STEP
C
      PRINT 7, (MP,N,X(MP,N), N=1,NT)
    7 FORMAT (1H ,7(15X,2HX(,I2,1H,,I2,3H) =,F13.4,/))
 1000 CONTINUE
 2000 STOP
      END
C     ********************************************************************
      SUBROUTINE PLANT
C
C     THIS SUBROUTINE EVALUATES THE OUTPUT ARRAY A(M,N) AND THE FIRST
C     PLANT DIFFERENTIAL DAD(M,I).
      DIMENSION X(101,21),A(101,21),DAD(101,21),ADA(101,21),XNORM(101),
     1PINGD(21),H(21)
     2,PLNT1(21),G(101,21),REVG(101,21),PLNT3(21)
      COMMON X,A,DAD,ADA,XNORM,PINGD,H,M,N,MT,NT,I,EPS,R,SW,XO
    1 FORMAT (1H ,4H M=I3,2X,2HI=I2,2X,7HA(M,I)=F12.3,2X,5HH(I)=F12.3,
     12X,9HDAD(M,I)=F12.3)
C
C     THE FOLLOWING STATEMENT SETS THE INITIAL PLANT OUTPUT
C
      A(M,1)=-EPS*(X(1,1)*X(1,1))
C
C     THE OPTIMIZATION INTERVAL T IS SUBDIVIDED INTO TWENTY EQUAL
C     INTERVALS AND A TRAPEZOIDAL RULE IS USED TO EVALUATE THE INTEGRAL
C     PART OF THE PLANT OUTPUT EXPRESSION
C
      DO 100 I=2,NT
      DO 200 JTAU=1,I
      IMTAU=I-JTAU
      BNT=NT-1
      ANT=(T+1.0)/BNT
  200 PLNT1(JTAU)=EXP(-IMTAU*ANT)*X(M,JTAU)
      IF (I-2) 210,210,220
  210 A(M,I)=0.5*PLNT1(1)*0.2+0.5*PLNT1(2)*0.2-EPS*(X(M,I)*X(M,I))
      GO TO 100
  220 SUM=0.0
      IM2=I-2
```

```
      DO 250 J=1,IM2
  250 SUM1=SUM1+PENT1(J+1)*0.2
      A(M,I)=0.5*PENT1(1)*0.2+SUM1+0.5*PENT1(I)*0.2-EPS*(X(M,I)*X(M,I))
  100 CONTINUE
C
C     THIS PART SETS THE VALUE OF G AND EVALUATES THE FIXED INPUT
C     VARIATION H
C
      DO 300 I=1,NT
  300 G(M,I)=1.0
      DO 305 I=1,NT
      REVG(M,I)=G(M,NT+1-I)
  305 H(I)=REVG(M,I)
C
C     THE PLANT DIFFERENTIAL IS EVALUATED FROM AN EXPLICIT RELATIONSHIP
C
      DAD(M,1)=-2.0*EPS*X(M,1)*H(1)
      DO 310 I=2,NT
      DO 320 JTAU=1,I
      IMTAU=I-JTAU
      BNT=NT-1
      ANT=(T+1.0)/BNT
  320 PENT3(JTAU)=EXP(-IMTAU*ANT)*H(JTAU)
      IF (I-2) 330,330,340
  330 DAD(M,I)=0.5*PENT3(1)*0.2+0.5*PENT3(2)*0.2-2.0*EPS*X(M,I)*H(I)
      GO TO 303
  340 SUM3=0.0
      IM2=I-2
      DO 350 J=1,IM2
  350 SUM3=SUM3+PENT3(J+1)*0.2
      DAD(M,I)=0.5*PENT3(1)*0.2+0.5*PENT3(I)*0.2+SUM3-2.0*EPS
     1*X(M,I)*H(I)
  303 PRINT 1, M,I,A(M,I),H(I),DAD(M,I)
  310 CONTINUE
      RETURN
      END
C     ******************************************************************
      SUBROUTINE ADJUNT
C
C     SUBROUTINE ADJUNT EVALUATES THE ADJOINT PLANT DIFFERENTIAL
C      USING THE REVERSE OPERATOR METHOD
C
      DIMENSION X(101,21),A(101,21),DAD(101,21),ADA(101,21),XNORM(101),
     1PINGO(21),H(21)
     2,REVDA(101,21),REVX(101,21)
      COMMON X,A,DAD,ADA,XNORM,PINGO,H,M,N,MT,NT,T,EPS,R,SW,XO
    1 FORMAT (1H ,15X,4HADA(,I3,1H,,I2,3H)=,E13.4)
      DO 100 I=1,NT
  100 REVDA(M,I)=DAD(M,NT+1-I)
      DO 200 I=1,NT
  200 REVX(M,I)=X(M,NT+1-I)
      DO 300 I=1,NT
      ADA(M,I)=REVDA(M,I)+2.0*EPS*(REVX(M,I)-X(M,I))
  300 PRINT 1, M,I,ADA(M,I)
      RETURN
      END
C     ******************************************************************
      SUBROUTINE NONCON
C
C     THIS SUBROUTINE ESTIMATES THE (M+1)ST ELEMENT OF THE CONTROL
```

```
C      SEQUENCE FOR THE UNCONSTRAINED PROBLEM VIA NEWTON'S MODIFIED
C      METHOD WITH A FIXED STEPSIZE
C
       DIMENSION X(101,21),A(101,21),BAD(101,21),ADA(101,21),XNORM(101),
      1PINGO(21),H(21)
       COMMON X,Z,BAD,ADA,XNORM,PINGO,H,M,N,MT,NT,I,EPS,R,SA,XO
       DO 50 I=1,NT
   50 X(M+1,I)=X(M,I)+ADA(M,I)*0.5
       RETURN
       END
C      ****************************************************************
       SUBROUTINE CON
C
C      SUBROUTINE CON ITERATIVELY GENERATES THE CONTROL SEQUENCE VIA
C      NEWTON'S MODIFIED METHOD FOR THE CONSTRAINED PROBLEM
C
       DIMENSION X(101,21),A(101,21),BAD(101,21),ADA(101,21),XNORM(101),
      1PINGO(21),H(21)
      2,PING(21)
       COMMON X,A,BAD,ADA,XNORM,PINGO,H,M,N,MT,NT,I,EPS,R,SA,XO
C
C      A VALUE FOR THE WEIGHTING PARAMETER ALAM IS CHOSEN BY TESTING
C      SO THAT THE CONTROL SIGNAL LIES ON THE BOUNDARY OF THE CONSTRAINT
C
       ALAM=0.001
C
C      THE NORM OF THE (M+1)ST ELEMENT OF THE CONTROL SEQUENCE, DENOTED
C      BY "AREA" IS EVALUATED AND COMPARED WITH THE CONSTANT CONSTRAINT
C      VALUE R.
C
   30 AREA=0.0
       DO 25 I=1,NT
       PING(I)=(X(M,I)-((2.0*ALAM*X(M,I)-ADA(M,I))/(2.0*(ALAM+
      1EPS))))*(X(M,I)-((2.0*ALAM*X(M,I)-ADA(M,I))/(2.0*(ALAM+
      2EPS))))
   25 AREA=AREA+PING(I)*0.2
       AREA=AREA-0.5*(PING(1)+PING(NT))*0.2
       IF (AREA-R) 10,20,20
   10 ALAM=ALAM+0.001
       IF (ALAM.LE.1.0) GO TO 30
       ALAM=0.0
       DO 5 I=1,NT
    5 X(M+1,I)=X(M,I)+ADA(M,I)*0.5
       RETURN
C
C      WITH THE CONTROL FUNCTION ON THE BOUNDARY OF THE CONSTRAINT THE
C      (M+1)ST ELEMENT OF THE SEQUENCE IS EVALUATED AND PRINTED
C
   20 DO 21 I=1,NT
   21 X(M+1,I)=X(M,I)-((2.0*ALAM*X(M,I)-ADA(M,I))/(2.0*(ALAM+
      1EPS)))
       PRINT 910, ALAM
  910 FORMAT (1H ,5X,'ALAM=',F5.3)
       RETURN
       END
C      ****************************************************************
```

```
C     THIS PROGRAM EVALUATES ITERATIVELY A SEQUENCE OF CONTROL FUNCTIONS
C     FOR A PLANT CONSISTING OF THE PARALLEL CONNECTION OF A LINEAR
C     INTEGRATOR AND A NONLINEAR PART.
C
C     THE ACCUMULATED PLANT OUTPUT IN A FIXED TIME INTERVAL IS MAXIMIZED.
C
C     THE MAGNITUDE OF THE CONTROL SIGNAL IS CONSTRAINED TO BE LESS OR
C     EQUAL TO A CONSTANT VALUE.
C
C     THE REVERSE OPERATOR METHOD IS UTILIZED FOR THE SOLUTION OF THE
C     IDENTIFICATION PROBLEM.
C
C     THE OPTIMAL INPUT SEQUENCE IS OBTAINED VIA NEWTON'S METHOD WITH
C     A FIXED STEPSIZE FOR THE UNCONSTRAINED PROBLEM. THE CONSTRAINED
C     SOLUTION DICTATES OPERATION ON THE BOUNDARY OF THE CONSTRAINT
C     (BANG-BANG TYPE) AND IS OBTAINED DIRECTLY.
C
C     THE OPTIMIZATION INTERVAL IS DIVIDED INTO FOUR EQUAL SUBINTERVALS.
C
C     VARIABLES USED IN THIS PROGRAM:
C
C     X     -CONTROL VARIABLE
C     XO    -OPTIMAL VALUE OF THE CONTROL SIGNAL CALCULATED THEORETICALLY
C     A     -PLANT OUTPUT
C     PINOD     -INTEGRAND OF PLANT OUTPUT EXPRESSION
C     PINOB     -INTEGRAND OF PLANT OUTPUT EXPRESSION WITH A FIXED INPUT
C     VARIATION
C     H     -FIXED INPUT VARIATION
C     G     -THE VARIATION OF THE ADJOINT DIFFERENTIAL EQUAL TO UNITY
C     DAD     -PLANT OUTPUT DIFFERENTIAL
C     ADA     -THE ADJOINT OF THE PLANT DIFFERENTIAL
C     XNORM     -THE NORM OF THE INPUT VARIABLE
C
C     CONSTANTS USED IN THIS PROGRAM:
C
C     MT     -TOTAL NUMBER OF ITERATION STEPS
C     NT     -NUMBER OF SAMPLING POINTS IN THE OPTIMIZATION INTERVAL
C     T     -OPTIMIZATION PERIOD
C     EPS     -WEIGHTING FACTOR OF THE NONLINEAR PART
C     R     -VALUE OF INPUT CONSTRAINT
C     ****************************************************************
C
C     THIS IS THE MAIN PROGRAM
C     IT SETS THE VALUES OF THE CONSTANT PARAMETERS, THE INITIAL
C     CONDITIONS, CALLS THE SUBROUTINES AND CHECKS FOR THE INPUT
C     CONSTRAINT.
C
      DIMENSION X(101,4),A(101,4),DAD(101,4),ADA(101,4),XNORM(101),
     1PINOD(4),XO(4),G(4),H(4)
      COMMON X,A,DAD,ADA,XNORM,PINOD,H,M,N,MT,NT,T,EPS,R,SG,XO
      READ 1, MT,NT
      READ 3, T,EPS
      READ 5, (X(1,N), N=1,NT)
      READ 5, (XO(N), N=1,NT)
      READ 5, (H(N), N=1,NT)
    1 FORMAT (3I10)
    3 FORMAT (3F10.4)
    5 FORMAT (4F10.4)
      SG=0.0
```

```
      DO 1000 I=1,MI
      CALL PLANT
      CALL ADJOINT
      CALL NORCON
      IF (SW-1.0) 500,2000,2000
C
C     THIS PART OF THE PROGRAM COMPARES THE ABSOLUTE VALUE OF THE
C     (M+1)ST ELEMENT OF THE CONTROL SEQUENCE WITH THE CONSTRAINT VALUE
C
  500 IF (ABS(X(M+1,1))-R(1)) 10,10,20
   10 IF (ABS(X(M+1,2))-R(2)) 11,11,20
   11 IF (ABS(X(M+1,3))-R(3)) 12,12,20
   12 IF (ABS(X(M+1,4))-R(4)) 800,800,20
   20 CALL CON
C
C     SUBROUTINE CON IS CALLED WHENEVER THE MAGNITUDE OF THE CONTROL
C     SIGNAL EXCEEDS THE CONSTRAINT VALUE R.
C
  800 MP=M+1
C
C     THE FOLLOWING STATEMENT PRINTS THE VALUES OF THE CONTROL FUNCTION
C     OBTAINED FROM THE PREVIOUS ITERATION STEP
C
      PRINT 7, (MP,N,X(MP,N), N=1,NI)
    7 FORMAT (1H ,/(15X,2HX(,I3,1H,,I2,2H)=,E13.4,/))
 1000 CONTINUE
 2000 STOP
      END
C     ****************************************************************
      SUBROUTINE PLANT
C
C     THIS SUBROUTINE EVALUATES THE OUTPUT ARRAY A(M,N) AND THE FIRST
C     PLANT DIFFERENTIAL DAD(M,I).
      DIMENSION X(101,4),A(101,4),DAD(101,4),ADA(101,4),XNORM(101),
     1PINGO(4),XO(4),R(4),H(4)
     2,PLNT1(4),G(101,4),EEVG(101,4),PLNT3(4)
      COMMON X,A,DAD,ADA,XNORM,PINGO,H,XO,N,MI,NI,T,EPS,R,SW,XU
    1 FORMAT (1H ,4H  M=I3,2X,2HI=I2,2X,7HA(M,I)=F12.3,2X,5HH(I)=F12.3,
     12X,9HDAD(M,I)=F12.3)
C
C     THE FOLLOWING STATEMENT SETS THE INITIAL PLANT OUTPUT
C
      A(M,1)=-EPS*(X(1,1)*X(1,1))
C
C     A TRAPEZOIDAL RULE IS USED TO EVALUATE THE INTEGRAL PART OF THE
C     PLANT OUTPUT EXPRESSION
C
      DO 100 I=2,NI
      DO 200 JTAU=1,I
      TMTAU=I-JTAU
  200 PLNT1(JTAU)=EXP(-TMTAU)*X(M,JTAU)
      IF (I-2) 210,210,220
  210 A(M,I)=0.5*PLNT1(1)+0.5*PLNT1(2)-EPS*(X(M,I)*X(M,I))
      GO TO 100
  220 SUM1=0.0
      IM2=I-2
      DO 250 J=1,IM2
  250 SUM1=SUM1+PLNT1(J+1)
      A(M,I)=0.5*PLNT1(1)+SUM1+0.5*PLNT1(I)-EPS*(X(M,I)*X(M,I))
  100 CONTINUE
```

```
C
C       THIS PART SETS THE VALUE OF G AND EVALUATES THE FIXED INPUT
C       VARIATION H
C
        DO 300 I=1,NT
  300   G(M,I)=1.0
        DO 305 I=1,NT
        REVG(M,I)=G(M,NT+1-I)
  305   H(I)=REVG(M,I)
C
C       THE PLANT DIFFERENTIAL IS EVALUATED FROM AN EXPLICIT RELATIONSHIP
C
        DAP(M,1)=-2.0*EPS*X(M,1)*H(1)
        DO 310 I=2,NT
        DO 320 JTAU=1,I
        IMTAU=I-JTAU
  320   PERT3(JTAU)=EXP(-IMTAU)*H(JTAU)
        IF (I-2) 330,330,340
  330   DAP(M,I)=0.5*PERT3(1)+0.5*PERT3(2)-2.0*EPS*X(M,I)*H(I)
        GO TO 303
  340   SUM=0.0
        IM2=I-2
        DO 350 J=1,IM2
  350   SUM=SUM+PERT3(J+1)
        DAP(M,I)=0.5*PERT3(1)+SUM+0.5*PERT3(I)-2.0*EPS*X(M,I)*H(I)
  303   PRINT 1, M,I,A(M,I),H(I),DAP(M,I)
  310   CONTINUE
        RETURN
        END
C       ******************************************************************
        SUBROUTINE ADJUNT
C
C       SUBROUTINE ADJUNT EVALUATES THE ADJOINT PLANT DIFFERENTIAL
C        USING THE REVERSE OPERATOR METHOD
C
        DIMENSION X(101,4),A(101,4),DAP(101,4),ADA(101,4),XNORM(101),
       1PINCO(4),XO(4),G(4),H(4)
       2,REVDA(101,4),REVX(101,4)
        COMMON X,A,DAP,ADA,XNORM,PINCO,H,M,N,MT,NT,I,EPS,R,SW,XO
    1   FORMAT (1H ,15X,4HADA(,I3,1H,,I2,3H)=,E13.4)
        DO 100 I=1,NT
  100   REVDA(M,I)=DAP(M,NT+1-I)
        DO 200 I=1,NT
  200   REVX(M,I)=X(M,NT+1-I)
        DO 300 I=1,NT
        ADA(M,I)=REVDA(M,I)+2.0*EPS*(REVX(M,I)-X(M,I))
  300   PRINT 1, M,I,ADA(M,I)
        RETURN
        END
C       ******************************************************************
        SUBROUTINE NONCON
C
C       THIS SUBROUTINE ESTIMATES THE (M+1)ST ELEMENT OF THE CONTROL
C       SEQUENCE FOR THE UNCONSTRAINED PROBLEM VIA NEWTON'S MODIFIED
C       METHOD WITH A FIXED STEPSIZE
C
        DIMENSION X(101,4),A(101,4),DAP(101,4),ADA(101,4),XNORM(101),
       1PINCO(4),XO(4),R(4),P(4)
       2,XP(101,4)
        COMMON X,A,DAP,ADA,XNORM,PINCO,H,M,N,MT,NT,I,EPS,R,SW,XO
```

```
      SW=1.0
      DO 10 I=1,NT
      XD(M,I)=ABS(XO(I)-X(M,I))
      IF (XD(M,I)-0.01) 10,20,20
   10 CONTINUE
      RETURN
   20 SW=0.0
      DO 50 I=1,NT
   50 X(M+1,I)=X(M,I)+ADA(M,I)*0.5
      RETURN
      END
C     ****************************************************************
      SUBROUTINE CON
C
C     SUBROUTINE CON GENERATES THE CONTROL FUNCTION DIRECTLY FROM
C     ITS OPTIMAL SOLUTION . THE CONSTRAINED PROBLEM IS CONSIDERED
C
      DIMENSION X(101,4),A(101,4),DAD(101,4),ADA(101,4),XNORM(101),
     1PREGD(4),XO(4),E(4),H(4)
      COMMON X,A,DAD,ADA,XNORM,PREGD,H,M,N,MT,NT,T,EPS,R,SW,XO
      DO 100 I=1,NT
  100 X(M+1,I)=E(I)*SGN(ADA(M,I))
      RETURN
      END
C     ****************************************************************
      FUNCTION SGN(Z)
C
C     THIS FUNCTION SUBPROGRAM EVALUATES THE SIGNUM OF A QUANTITY Z
C     IF Z IS POSITIVE THE SIGNUM IS +1
C     IF Z IS NEGATIVE THE SIGNUM IS -1
C     IF Z IS ZERO THE SIGNUM IS TAKEN TO BE ZERO
      IF (Z) 1,2,3
    1 SGN=-1.0
    2 SGN=0.0
    3 SGN=1.0
      RETURN
      END
C     ****************************************************************
```

```
C     THIS PROGRAM EVALUATES ITERATIVELY A SEQUENCE OF CONTROL FUNCTIONS
C     FOR A CHEMICAL REACTOR PROCESS.
C
C     THE PLANT IS DESCRIBED BY A SET OF ORDINARY DIFFERENTIAL EQUATIONS.
C
C     THE REACTOR TEMPERATURE IS TAKEN AS THE CONTROL VARIABLE.
C
C     THE TOTAL OUTPUT PRODUCT IN A FIXED TIME INTERVAL IS MAXIMIZED.
C
C     THE MAGNITUDE OF THE CONTROL SIGNAL IS CONSTRAINED TO BE LESS OR
C     EQUAL TO A CONSTANT VALUE.
C
C     METHOD:
C
C     A FOURTH ORDER RUNGE-KUTTA INTEGRATION PROCESS IS USED FOR THE
C     NUMERICAL SOLUTION OF THE DIFFERENTIAL EQUATIONS.
C
C     AN APPROXIMATE ESTIMATE OF THE PLANT OUTPUT DIFFERENTIAL IS
C     OBTAINED BY PERTURBING THE INPUT FUNCTION.
C
C     THE ADJOINT DIFFERENTIAL IS OBTAINED VIA THE REVERSE OPERATION
C     METHOD.
C
C     THE OPTIMAL INPUT SEQUENCE IS GENERATED UTILIZING NEWTON'S
C     MODIFIED SCHEME WITH A FIXED STEPSIZE FOR THE UNCONSTRAINED
C     PROBLEM.
C
C     THE CONSTRAINED SOLUTION IS OBTAINED DIRECTLY.
C
C     THE OPTIMIZATION INTERVAL IS DIVIDED INTO SIX EQUAL SUBINTERVALS.
C
C
C     VARIABLES USED IN THIS PROGRAM:
C
C     X     -CONTROL VARIABLE
C     A     -PLANT OUTPUT
C     AINC     -PLANT OUTPUT WITH A FIXED INPUT VARIATION
C     H     -FIXED STEPSIZE FOR THE RUNGE-KUTTA PROCESS
C     DA     -PLANT OUTPUT DIFFERENTIAL
C     ADA     -THE ADJOINT OF THE PLANT DIFFERENTIAL
C     VEC1     -RESULTANT PLANT OUTPUT VALUE USING THE RUNGE-KUTTA METHOD
C     VEC2     -RESULTANT PLANT OUTPUT VALUE WITH A FIXED INPUT VARIATION
C     PFY     -Y-COEFFICIENT OF THE LINEARIZED VARIATIONAL PLANT EQUATION
C     PFX     -X-COEFFICIENT OF THE LINEARIZED VARIATIONAL PLANT EQUATION
C
C
C     CONSTANTS USED IN THIS PROGRAM:
C
C     MT     -TOTAL NUMBER OF ITERATIVE STEPS
C     NT     -NUMBER OF SAMPLING POINTS IN THE OPTIMIZATION INTERVAL
C     GAMMA     -FIXED INPUT INCREMENT
C     R     -VALUE OF INPUT CONSTRAINT
C
C***********************************************************************
C     THIS IS THE MAIN PROGRAM
C
C     IT READS THE VALUES OF THE CONSTANT PARAMETERS, SETS THE INITIAL
C     CONDITIONS, CALLS THE SUBROUTINES AND CHECKS FOR THE INPUT
C     CONSTRAINT
```

```
C
      DIMENSION X(101,6),A(101,6),DA(101,6),ADA(101,6)
      COMMON X,A,DA,ADA,M,N,MT,NT,R,SW,GAMMA
      READ 1, MT,NT
      READ 3, GAMMA,R
      READ 5, X(1,1),A(1,1)
    1 FORMAT (2I10)
    3 FORMAT (2F10.4)
    5 FORMAT (2F10.4)
      SW=0.0
      DO 1000 M=1,MT
      CALL PLANT
      CALL ADJOINT
      CALL GRADN
C
C     THIS PART OF THE PROGRAM COMPARES THE ABSOLUTE VALUE OF THE
C     (M+1)ST ELEMENT OF THE CONTROL VECTOR WITH THE VALUE OF THE
C     INPUT CONSTRAINT
C
      IF (ABS(X(M+1,1))-R) 10,10,20
   10 IF (ABS(X(M+1,2))-R) 11,11,20
   11 IF (ABS(X(M+1,3))-R) 12,12,20
   12 IF (ABS(X(M+1,4))-R) 13,13,20
   13 IF (ABS(X(M+1,5))-R) 14,14,20
   14 IF (ABS(X(M+1,6))-R) 30,30,20
   20 CALL CON
C
C     SUBROUTINE CON IS CALLED WHEN THE MAGNITUDE OF THE CONTROL
C     FUNCTION EXCEEDS THE VALUE OF THE CONSTRAINT
C
C
C     THE FOLLOWING PART COMPARES THE PRESENT LEVEL OF THE CONTROL
C     SIGNAL WITH THE PREVIOUS ONE. IF THEY ARE IDENTICAL THE PROGRAM
C     IS TERMINATED
C
      IF (X(M+1,1)-X(M,1)) 800,30,800
   30 IF (X(M+1,2)-X(M,2)) 800,31,800
   31 IF (X(M+1,3)-X(M,3)) 800,32,800
   32 IF (X(M+1,4)-X(M,4)) 800,33,800
   33 IF (X(M+1,5)-X(M,5)) 800,34,800
   34 IF (X(M+1,6)-X(M,6)) 800,35,800
   35 SW=1.0
      PRINT 9, SW
    9 FORMAT (1H ,5X,'SW= ',F3.1)
      GO TO 2000
  800 MP=M+1
C
C     THE VALUES OF THE CONTROL FUNCTION OBTAINED FROM THE PREVIOUS
C     ITERATION ARE PRINTED
C
      PRINT 7, (MP,N,X(MP,N), N=1,NT)
    7 FORMAT (1H ,/(15X,2HX(,I3,1H,,I2,3H)= ,E16.8,/))
 1000 CONTINUE
 2000 STOP
      END
C     *****************************************************************
      SUBROUTINE PLANT
C
C     THIS SUBROUTINE EVALUATES THE OUTPUT ARRAY A(M,N) AND THE FIRST
C     PLANT DIFFERENTIAL DA(M,N)
```

```
C
C
      DIMENSION X(101,6),A(101,6),DA(101,6),ADA(101,6)
     1,VEC1(101,6),VEC2(101,6),AINC(101,6)
     2,PEY(101,6),PEX(101,6)
      COMMON X,A,DA,ADA,M,N,MT,NT,R,SW,GAMMA
      IF (M.NE.1) GO TO 1
      CALL PLANT1
C
C     SUBROUTINE PLANT1 EVALUATES THE INITIAL VALUE OF X
C
C
C     THE FOLLOWING STATEMENTS DETERMINE THE OUTPUT ARRAY A(M,I) USING
C     A FOURTH ORDER RUNGE-KUTTA PROCESS
C
C     FUN1 IS A TWO ARGUMENT FUNCTION SUBPROGRAM. DY/DT   FUN1(Y,Z)
C
    1 YI=0.75-A(M,1)
      H=0.5
      DO 2 I=1,NT
      T1=H*FUN1(Y,X(M,I))
      T2=H*FUN1(Y+T1/2.,X(M,I))
      T3=H*FUN1(Y+T2/2.,X(M,I))
      T4=H*FUN1(Y+T3,X(M,I))
      Y=Y+(T1+2.*T2+2.*T3+T4)/6.
      VEC1(M,I)=Y
      A(M,I)=0.75-VEC1(M,I)
    2 PRINT 5, M,I,A(M,I)
    5 FORMAT (1H ,4H  M=,I3,2X,2HI=,I1,2X,7HA(M,I)=,E16.8)
C
C
C     THIS PART EVALUATES THE OUTPUT ARRAY WHEN THE INPUT IS PERTURBED
C     BY A FIXED VARIATION
C
      YI=0.75-A(M,1)
      H=0.5
      Y=YI
      DO 4 I=1,NT
      T1=H*FUN1(Y,X(M,I)+GAMMA)
      T2=H*FUN1(Y+T1/2.,X(M,I)+GAMMA)
      T3=H*FUN1(Y+T2/2.,X(M,I)+GAMMA)
      T4=H*FUN1(Y+T3,X(M,I)+GAMMA)
      Y=Y+(T1+2.*T2+2.*T3+T4)/6.
      VEC2(M,I)=Y
      AINC(M,I)=0.75-VEC2(M,I)
C
C     AN APPROXIMATE ESTIMATE OF THE PLANT DIFFERENTIAL IS OBTAINED
C
      DA(M,I)=(AINC(M,I)-A(M,I))/GAMMA
    4 PRINT 3, M,I,DA(M,I)
    3 FORMAT (1H ,4H  M=,I3,2X,2HI=,I1,2X,8HDA(M,I)=,E16.8)
C
C     THE VALUES OF THE COEFFICIENTS OF THE LINEARIZED VARIATIONAL
C     OUTPUT EQUATION ARE DETERMINED
C
      DO 10 I=1,NT
      PEY(M,I)=-1.0-EXP(-1.0/X(M,I))
      PEX(M,I)=-1.-(0.75-A(M,I))*EXP(-1./X(M,I))/(X(M,I)*X(M,I))
   10 PRINT 9, M,I,PEY(M,I),PEX(M,I)
    9 FORMAT (1H ,4H  M=,I3,2X,2HI=,I1,2X,9HPEY(M,I)=,E16.8,5X
```

```
     1,9HPLX(M,I)=,E16.8)
      RETURN
      END
C     ************************************************************************
      SUBROUTINE PLANTI
C
C
C     THIS SUBROUTINE EVALUATES THE INITIAL VALUE OF X
C
      DIMENSION X(101,6),A(101,6),DA(101,6),ADA(101,6)
      COMMON X,A,DA,ADA,M,N,NT,NI,H,SG,GAMMA
      Y1=0.75
      Z1=X(1,1)
      H=0.5
      Y=Y1
      Z=Z1
      DO 2 I=1,NT
C
C     FUN1 IS A TWO ARGUMENT FUNCTION SUBPROGRAM. DY/DT = FUN1(Y,Z)
C     FUN2 IS A TWO ARGUMENT FUNCTION SUBPROGRAM. DZ/DT = FUN2(Y,Z)
C
      T1=H*FUN1(Y,Z)
      P1=H*FUN2(Y,Z)
      T2=H*FUN1(Y+T1/2.,Z+P1/2.)
      P2=H*FUN2(Y+T1/2.,Z+P1/2.)
      T3=H*FUN1(Y+T2/2.,Z+P2/2.)
      P3=H*FUN2(Y+T2/2.,Z+P2/2.)
      T4=H*FUN1(Y+T3,Z+P3)
      P4=H*FUN2(Y+T3,Z+P3)
      Y=Y+(T1+2.*T2+2.*T3+T4)/6.
      Z=Z+(P1+2.*P2+2.*P3+P4)/6.
      A(M,I)=0.75-Y
      X(M,I)=Z
      PRINT 5,M,I,A(M,I)
    2 PRINT 7,M,I,X(M,I)
    5 FORMAT (1H ,4H M =,I3,2X,2H I=,I1,3X,7HA(M,I)=,E16.8)
    7 FORMAT (1H ,4H M =,I3,2X,2H I=,I1,3X,7HX(M,I)=,E16.8)
      RETURN
      END
C     ************************************************************************
      FUNCTION FUN1(Y,Z)
C
C
C     THIS IS A TWO ARGUMENT FUNCTION SUBPROGRAM.  DY/DT = FUN1(Y,Z)
C
      FUN1=1.-Y*(1.+EXP(-1./Z))
      RETURN
      END
C
C     ************************************************************************
      FUNCTION FUN2(Y,Z)
C
C
C     THIS IS A TWO ARGUMENT FUNCTION SUBPROGRAM.  DZ/DT = FUN2(Y,Z)
C
      FUN2=(1.-Z)+Y*EXP(-1./Z)
      RETURN
      END
C
C     ************************************************************************
```

```
      SUBROUTINE ADJUNT
C
C
C     SUBROUTINE ADJUNT EVALUATES THE ADJOINT PLANT DIFFERENTIAL
C      USING THE REVERSE OPERATOR METHOD
C
      DIMENSION X(101,6),A(101,6),DA(101,6),ADA(101,6)
     1,REVDA(101,6)
      COMMON X,A,DA,ADA,M,N,MI,NI,R,SP,GAMMA
      DO 100 I=1,NI
  100 REVDA(M,I)=DA(M,NI+1-I)
      DO 200 I=1,NI
      ADA(M,I)=REVDA(M,I)
  200 PRINT 1, M,I,ADA(M,I)
    1 FORMAT (1H ,15X,4HADA(,I3,1H,,I1,2H) ,E16.8)
      RETURN
      END
C
C     ******************************************************************
      SUBROUTINE NEWCON
C
C
C     THIS SUBROUTINE GENERATES THE (M+1)ST ITERATE OF THE CONTROL
C     SEQUENCE FOR THE UNCONSTRAINED PROBLEM VIA NEWTON'S MODIFIED
C     METHOD WITH A FIXED STEPSIZE
C
      DIMENSION X(101,6),A(101,6),DA(101,6),ADA(101,6)
      COMMON X,A,DA,ADA,M,N,MI,NI,R,SP,GAMMA
      DO 50 I=1,NI
   50 X(M+1,I)=X(M,I)+ADA(M,I)*0.1
      RETURN
      END
C
C     ******************************************************************
      SUBROUTINE CON
C
C
C     SUBROUTINE CON GENERATES THE CONTROL FUNCTION DIRECTLY FROM
C     ITS OPTIMAL SOLUTION . THE CONSTRAINED PROBLEM IS CONSIDERED
C
      DIMENSION X(101,6),A(101,6),DA(101,6),ADA(101,6)
      COMMON X,A,DA,ADA,M,N,MI,NI,R,SP,GAMMA
      DO 100 I=1,NI
  100 X(M+1,I)=R*SGN(ADA(M,I))
      RETURN
      END
C
C     ******************************************************************
      FUNCTION SGN(Z)
C
C
C     THIS FUNCTION SUBPROGRAM EVALUATES THE SIGNUM OF A QUANTITY Z
C
C     IF Z IS POSITIVE THE SIGNUM IS +1
C     IF Z IS NEGATIVE THE SIGNUM IS -1
   IF Z IS ZERO THE SIGNUM IS TAKEN TO BE ZERO
C
      IF (Z) 1,2,3
    1 SGN=-1.0
    2 SGN=0.0
```

```
    3 SGN=1.0
      RETURN
      END
C
```